Science Policy Study—Hearings Volume 22
RESEARCH AND PUBLICATIONS PRACTICES

HEARING

BEFORE THE

TASK FORCE ON SCIENCE POLICY

OF THE

COMMITTEE ON SCIENCE AND TECHNOLOGY HOUSE OF REPRESENTATIVES

NINETY-NINTH CONGRESS

SECOND SESSION

MAY 14, 1986

[No. 165]

Printed for the use of the
Committee on Science and Technology

U.S. GOVERNMENT PRINTING OFFICE

63-471 WASHINGTON : 1987

For sale by the Superintendent of Documents, Congressional Sales Office
U.S. Government Printing Office, Washington, DC 20402

H701-44

COMMITTEE ON SCIENCE AND TECHNOLOGY

DON FUQUA, Florida, *Chairman*

ROBERT A. ROE, New Jersey
GEORGE E. BROWN, Jr., California
JAMES H. SCHEUER, New York
MARILYN LLOYD, Tennessee
TIMOTHY E. WIRTH, Colorado
DOUG WALGREN, Pennsylvania
DAN GLICKMAN, Kansas
ROBERT A. YOUNG, Missouri
HAROLD L. VOLKMER, Missouri
BILL NELSON, Florida
STAN LUNDINE, New York
RALPH M. HALL, Texas
DAVE McCURDY, Oklahoma
NORMAN Y. MINETA, California
MICHAEL A. ANDREWS, Texas
BUDDY MacKAY, Florida**
TIM VALENTINE, North Carolina
HARRY M. REID, Nevada
ROBERT G. TORRICELLI, New Jersey
RICK BOUCHER, Virginia
TERRY BRUCE, Illinois
RICHARD H. STALLINGS, Idaho
BART GORDON, Tennessee
JAMES A. TRAFICANT, Jr., Ohio

MANUEL LUJAN, Jr., New Mexico*
ROBERT S. WALKER, Pennsylvania
F. JAMES SENSENBRENNER, Jr., Wisconsin
CLAUDINE SCHNEIDER, Rhode Island
SHERWOOD L. BOEHLERT, New York
TOM LEWIS, Florida
DON RITTER, Pennsylvania
SID W. MORRISON, Washington
RON PACKARD, California
JAN MEYERS, Kansas
ROBERT C. SMITH, New Hampshire
PAUL B. HENRY, Michigan
HARRIS W. FAWELL, Illinois
WILLIAM W. COBEY, Jr., North Carolina
JOE BARTON, Texas
D. FRENCH SLAUGHTER, Jr., Virginia
DAVID S. MONSON, Utah

HAROLD P. HANSON, *Executive Director*
ROBERT C. KETCHAM, *General Counsel*
REGINA A. DAVIS, *Chief Clerk*
JOYCE GROSS FREIWALD, *Republican Staff Director*

SCIENCE POLICY TASK FORCE

DON FUQUA, Florida, *Chairman*

GEORGE E. BROWN, Jr., California
TIMOTHY E. WIRTH, Colorado
DOUG WALGREN, Pennsylvania
HAROLD L. VOLKMER, Missouri
STAN LUNDINE, New York
NORMAN Y. MINETA, California
HARRY M. REID, Nevada
RICK BOUCHER, Virginia
RICHARD H. STALLINGS, Idaho

MANUEL LUJAN, Jr., New Mexico*
TOM LEWIS, Florida
DON RITTER, Pennsylvania
SID W. MORRISON, Washington
RON PACKARD, California
JAN MEYERS, Kansas
HARRIS W. FAWELL, Illinois
D. FRENCH SLAUGHTER, Jr., Virginia

JOHN D. HOLMFELD, *Study Director*
R. THOMAS WEIMER, *Republican Staff Member*

---

*Ranking Republican Member.
**Serving on Committee on the Budget for 99th Congress.

CONTENTS

WITNESSES

	Page
May 14, 1986:	
Dr. Sandra Penum, guest scholar, Center for Public Policy Education, the Brookings Institution, Washington, DC	2
Prepared statement	6
Discussion	12
Walter W. Stewart, Research Physicist, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD; and Dr. Ned Feder, Chief, Section on Biophysical Histology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD	17
Prepared statement	27
Walter W. Stewart and Ned Feder, "Professional Practices Among Biomedical Scientists: A Study of a Sample Generated by an Unusual Event," draft of an unpublished manuscript, May 11, 1986	56
Discussion	136
Dr. Patricia Woolf, lecturer, Sociology Department, Princeton University, Princeton, NJ	141
Prepared statement	147
Discussion	155
Appendix 1: Statement for the record by: Bancroft Littlefield, Jr., Esquire, Foley, Hoag and Eliot, Boston, MA, and response by Walter W. Stewart and Ned Feder, National Institutes of Health, Bethesda, MD	159
Appendix 2: Article by W. Stewart and N. Feder and comment by E. Brunwald in Nature magazine	202
Walter W. Stewart and Ned Feder, "The Integrity of the Scientific Literature," Nature, vol. 325, January 15, 1987, pp. 207-214	203
Dr. Eugene Braunwald, "On Analysing Scientific Fraud," Natue, vol. 325, January 15, 1987, pp. 215-216	211
Appendix 3: Press comments on Stewart and Feder Paper	213
Daniel E. Koshland, Jr., "Fraud in Science," Science, vol. 215, No. 4785, January 9, 1987	214
"Fraud, Libel, and the Literature," Nature, vol. 235, January 15, 1987, pp. 181-182	215
Barbara J. Culliton, "Integrity of Research Papers Questioned," Science, January 23, 1987, pp. 422-423	217
Boyce Rensberger, "Fraud, Laxity in Research Are Detailed," Washington Post, January 27, 1987	219
Andy Coghlan, "Researchers Roll Back the Frontiers of Fraud," New Scientist, January 22, 1987	220
Nicholas Wade, "Fraud and Garbage in Science," the New York Times, January 29, 1987	221
Arnold S. Relman, "Honest Science Errors Aren't Garbage," the New York Times, February 7, 1987	221
David L. Wheeler, "Study Alleging Scientific Misconduct Is Published; Findings Disputed," the Chronicle of High Education, February 1987	222
Sharon Begley, Mary Heger, and Shawn Doherty, "Tempests in a Test Tube," Newsweek, February 2, 1987, p. 64	223
Appendix 4: Reports on the Darsee Case from Emory and Harvard Universities and the National Institutes of Health	224
"Report of Ad Hoc Committee To Evaluate Research of Dr. John R. Darsee at Emory University," pp. 276-305	225

Appendix 4—Continued Page

"Medical School Dean Releases Report on Falsification of Research Data," Harvard Gazette, January 29, 1982 255

"Report of the National Heart, Lung, and Blood Institute Special Panel to Review Alleged Misconduct at Brigham and Women's Hospital/Harvard Medical School, Boston, Massachusetts," June 29, 1982 262

Memorandum from the Associate Director for Extramural Research and Training, National Institutes of Health to the Director, National Institutes of Health, subject: Investigation of alleged misconduct at Brigham and Women's Hospital and proposed actions—Decision, January 5, 1983 301

HHS News, for release February 16, 1983 312

Comments on the NIH Report 314

Memorandum from Eugene Braunwald, M.D., chairman, Department of Medicine, physician-in-chief, Brigham and Beth Israel Hospitals, to members of the committee on senior NIH staff convened by Dr. Raub to make recommendations to Dr. Wyngaarden concerning the report of the National Heart, Lung, and Blood Institute special panel to review alleged misconduct at Brigham and Women's Hospital/Harvard Medical School, Boston, MA, subject: Comments concerning NHLBI panel's report and range of possible NIH actions, September 21, 1982 315

Letter from Daniel C. Tosteson, dean, Harvard Medical School, to William F. Raub, Ph.D., Associate Director for Extramural Research and Training, National Institutes of Health, dated September 24, 1982 399

Letter from Robert A. Kloner, M.D., Ph.D. assistant professor of medicine, Harvard Medical School, to Dr. William F. Raub, Associate Director for Extramural Research and Training, National Insitutes of Health, dated September 22, 1982 401

Memorandum from Robert A. Kloner, M.D., Ph.D., to Dr. William F. Raub, Associate Director for Extramural Research and Training, National Institutes of Health, subject: Report of the NHLBI special panel to review alleged misconduct at the Brigham and Women's Hospital/Harvard Medical School, dated September 14, 1982 404

Memorandum from Eugene Braunwald, M.D., chairman, Department of Medicine, physician-in-chief, Brigham and Beth Israel Hospitals, to Dr. William F. Raub, Associate Director for Extramural Research and Training, National Institutes of Health, subject: Scientific misconduct of John R. Darsee, M.D., December 13, 1982 425

Letter from Eugene Braunwald, M.D., chairman, Department of Medicine, physician-in-chief, Brigham and Beth Israel Hospitals, to Dr. William F. Raub, Associate Director for Extramural Research and Training, National Institutes of Health, February 11, 1983 437

Appendix 5: Press comments on Darsee Case 439

William J. Broad, "Harvard Delays in Reporting Fraud," Science, vol. 215, January 29, 1982, pp. 478-482 440

William J. Broad, "Report Absolves Harvard in Case of Fakery," Science, vol. 215, February 12, 1982, pp. 874-876 445

Barbara J. Culliton, "Coping With Fraud: The Darsee Case," Science, vol., 220, April 1, 1983, pp. 31-34 448

Richard A. Knox, "The Harvard Fraud Case: Where Does the Problem Lie?", JAMA, vol. 249, No. 14, April 8, 1983, pp. 1797-1807 452

Barbara J. Culliton, "Emory Reports on Darsee's Fraud," Science, vol. 220, May 27, 1983, p. 936 461

Donald O. Nutter, Steven B. Haymsfield, and James F. Glenn, "RETRACTION, Darsee, JR, Heymsfield SB, Nutter, DO. Dypertrophic cardiomyopathy and human leukocyte antigen linkage: differentiation of two forms of hypertrohic cardiomyopathy. N Engl J Med 1979; 300:877-82," the New England Journal of Medicine, vol. 308, No. 23, June 9, 1983, p. 1400 462

Steven B. Heymsfield, and James F. Glenn, "RETRACTION, Daree JR, Heymsfield SB. Decreased myocardial taurine levels and hypertaurinuria in a kindred with mitral-valve prolapse and congestive cardiomyopathy. N Engl J Med 1981; 304:129-35," the New England Journal of Medicine, vol. 308, No. 23, June 9, 1983, p. 1400 462

Arnold S. Relman, M.D., "Lessons From the Darsee Affair," the New England Journal of Medicine, vol. 308, No. 23, June 9, 1983, pp. 1415-1417 463

Richard A. Knox, "Deeper Problems for Darsee: Emory Probe," JAMA, vol. 249, No. 21, June 3, 1983, pp. 2867-2873, 2876 466

Page

Appendix 6: Overviews of misconduct and fraud in research 471
Daryl E. Chubin, "Misconduct in Research: An Issue of Science Policy and Practice," Minerva, Summer 1985, vol. 23, No. 2, pp. 175-202 472
Marcia Angell, "Fraud in Science," Science, vol. 219, March 25, 1983 500
Walter Gratzer, "Trouble at t'lab," Nature, vol. 302, April 28, 1983 503
Penelope J. Greene, Jane S. Durch, Wendy Horwitz, and Vlwyn S. Hooper, "Policies for Responding to Allegations of Fraud in Research," Minerva, Summer 1985, vol. 23, No. 2, pp. 204-215 505
"Policy for Dealing With Faculty Fraud in Research," University of Florida, Minerva, Summer 1985, vol. 23, No. 2, pp. 175-202 518
Appendix 7: General discussion of research publication practices, ethics and fraud:
Ward Pigman and Emmett B. Carmichael, "An Ethical Code for Scientists," Science, vol. 111, June 16, 1950, pp. 643-647 523
Philip W. Majerus, "Fraud in Medical Research: Presidential Address Delivered Before the 74th Annual Meeting of the American Society for Clinical Investigation, Washington, D.C., May 8, 1982," J. Clin. Invest., vol. 70, July 1982, pp. 213-217 528
"Authorship From the Reader's Side," Annals of Internal Medicine, vol. 97, No. 4, October 1982, pp. 613-614 533
Lt. Col. Kenneth D. Burman, " 'Hanging From the Masthead': Reflections on Authorship," Annals of Internal Medicine, vol. 97, No. 4, October 1982, pp. 602-605 535
C. Ian Jackson and John W. Prados, "Honor in Science," American Scientist, vol. 71, September-October 1983, pp. 462-464 539
DeWitt Stetten, Jr., "Reported Laboratory Frauds in Biomedical Sciences," Science, vol. 226, December 21, 1984 542
Thomas P. Stossel, M.D., "SPEED: An Essay on Biomedical Communication," the New England Journal of Medicine, July 11, 1985, vol. 313, No. 2 544
Marcia Angell, M.D., "Publish or Perish: A Proposal," Annals of Internal Medicine, vol. 104, No. 2, February 1986, pp. 261-262 548
John C. Bailer III, "Science, Statistics, and Deception," Annals of Internal Medicine, vol. 104, No. 2, February 1986, pp. 259-260 550
Patricia K. Woolf, "Pressure to Publish and Fraud In Science," Annals of Internal Medicine, vol. 104, No. 2, February 1986, pp. 254-256 552
Robert G. Petersdorf, "The Pathogenesis of Fraud in Medical Science," Annals of Internal Medicine, vol. 104, No. 2, February 1986, pp. 252-254 555
Edward J. Huth, "Irresponsible Authorship and Wasteful Publication," Annals of Internal Medicine, vol. 104, No. 2, February 1986, pp. 257-259 558
Edward J. Huth, "Guidelines on Authorship of Medical Paprs," Annals of Internal Medicine, vol. 104, No. 2, February 1986, pp. 269-274 561
Edward J. Huth, "Abuses and Uses of Authorship," Annals of Internal Medicine, vol. 104, No. 2, February 1986, pp. 266-267 567
Appendix 8: Proposed National Science Foundation regulation for dealing with misconduct in science and engineering research, Federal Register, vol. 52, vol. 7, February 10, 1987, pp. 4158-4161 569
Appendix 9: Additional materials for the record:
"Report of the Association of American Universities Committee on the Integrity of Research," approved by the Joint Committee on Health Policy of the American Association of American Universities, the American Council on Education, and the National Association of State Universities and Land-Grant Colleges, 1983 574
Donald Kennedy, "On Academic Authorship," September 1985 580

RESEARCH AND PUBLICATIONS PRACTICES

WEDNESDAY, MAY 14, 1986

HOUSE OF REPRESENTATIVES,
COMMITTEE ON SCIENCE AND TECHNOLOGY,
TASK FORCE ON SCIENCE POLICY,
Washington, DC.

The task force met, pursuant to call, at 10:05 a.m., in room 2318, Rayburn House Office Building, Hon. Don Fuqua (chairman of the task force) presiding.

Mr. FUQUA. This is the final and concluding day of our task force hearings on scientific research in the mission agencies.

The first phase of these hearings took place last fall when we surveyed the research policies of most of the mission agencies including USDA, NASA, EPA, the Department of Energy, and others.

Last month our hearings on this subject were completed with 3 days of hearings on the future of research in the three most important biomedical research agencies: the National Institutes of Health, Veterans' Administration, and Food and Drug Administration.

The highly favorable picture which emerged from those recent hearings may, however, only represent one side of the coin. No one questions the great strides which biomedical research has made in the last 40 years, but there is also a more troubling side to that picture. A small but growing number of issues concerning the way in which the biomedical research is managed and operates suggest that there may be trouble ahead if attention is not paid now.

In our hearing today we will focus on several aspects of that question. We will attempt to determine if the present system of incentives for the performance of research have side effects that are undesirable and which seriously distract from the primary task of finding and using new scientific knowledge.

We are grateful to the scientists who have agreed to come before us today and discuss these difficult questions. And I emphasize that our purpose on this occasion is not to find fault with any individual or institution. Our purpose is to look ahead and to find new pathways that will benefit American science and the American taxpayer in future years.

We are pleased for our first witness to be Dr. Sandra Panem, Center for Public Policy Education at the Brookings Institution here in Washington.

Dr. Panem, we are very pleased to have you here this morning.

DR. SANDRA PANEM

Sandra Panem is currently a Guest Scholar in the Center for Public Policy Education at the Brookings Institution. She received a B.S. degree in biochemistry (1966) and a Ph.D. degree in microbiology (1970) from the University of Chicago.

Dr. Panem has been Science Advisor and Director of the Science Group in the Office of Policy, Planning and Evaluation, U.S. Environmental Protection Agency (1983–1985). Prior to joining EPA, she was a Science and Public Policy Fellow at the Brookings Institution (1982–1983).

As an experimental virologist, Dr. Panem has been Associate Professor of Pathology and Laboratory Medicine at Emory University, Atlanta; a member of the faculty of the University of Chicago; and a visting professor at the Universitaet Zurich. Her interest as an experimental scientist concerned the role of viral infection in human leukemia and autoimmune disease. Dr. Panem has received fellowship and career development awards from the Leukemia Society of America, the Damon Runyon Fund for Cancer Research and the National Institutes of Health. She is the author of numerous research publications and has also written about biomedical research policy for the lay and scientific press.

Her book, *The Interferon Crusade,* examines the interferon field as a case study of the biomedical research community and the developing biotechnological industry. In addition, she has edited *Science, Public Policy and Environmental Risk;* and *Biotechnology: Implications for Public Policy.* She is now completing a study on the crisis management of medical research for the Twentieth Century Fund.

In association with her work in science and public policy, Dr. Panem has been a Kellogg National Fellow, and a Science and Public Policy Fellow in the economics studies program at the Brookings Institution.

STATEMENT OF DR. SANDRA PANEM, GUEST SCHOLAR IN THE CENTER FOR PUBLIC POLICY EDUCATION, THE BROOKINGS INSTITUTION, WASHINGTON, DC

Dr. PANEM. I am Sandra Panem, a guest scholar at the Brookings Institution. I am pleased to be able to appear before this committee and to participate in your discussions concerning the role of the Federal Government in conducting and supporting basic and applied biomedical research. I commend the committee for convening these hearings. They can only increase appreciation of the importance of Federal support for biomedical research, as well as the need for accountability by the scientific community in response to that support.

I have been asked to comment on issues affecting the capability of the life sciences research community to function in an optimum fashion. In anticipation of this hearing, therefore, I have examined the testimony of many distinguished members of the community who appeared earlier this month. Two central issues correctly emerge from that diverse testimony. The first is whether current levels of Federal support for biomedical research are sufficient to take maximum advantage of the potential of contemporary science. The second is whether current policies allow the most efficient expenditure of resources for biomedical research, regardless of the level of funding.

I need only briefly comment on these issues as I am in general agreement with the majority of the other witnesses. The technological advances of the genetic revolution, now more than a decade old, have provided biomedical research opportunities that have not yet begun to be tapped. The limiting factors for progress in research are adequate numbers of trained personnel, sufficient laboratory facilities and predictable and continuous funding. Thus the limiting factors are not scientific, but are issues of social choice. Such questions are eminently suited to resolution in the Congress.

There can be no question that biomedical science can profitably utilize increasing resources in the foreseeable future for training, for maintaining capital facilities, and for research.

The second central issue is how to foster efficient use of any level of funding. Here, new public policy is needed to foster efficiency by providing consistent and predictable funding. This will allow better planning and eliminate the inefficiencies of disrupting ongoing research before results can be gathered, as well as the inefficiency of time lost in stopping and then starting up research programs. Implementing policies which would allow a multiyear budget cycle for the National Institutes of Health and the Centers for Disease Control would help achieve the goal of more efficient use of resources.

As I read the committee report which lays out the mandate for this congressional Science Policy Task Force, several underlying assumptions emerge. There is a belief that investment in basic biomedical research is justified because improvements in health derive from research findings. As basic research is therefore in society's interest, it is appropriate for the Federal Government to support basic research within Government facilities, as well as within universities, research institutes, and other organizations. These ideas have guided Federal support of biomedical research since the 1930's. Thus, we have traditionally supported research as an investment for the future. I would endorse each of these ideas.

Yet there is an additional and a powerful reason to vigilantly support the basic research enterprise that I believe has not received sufficient emphasis. A strong basic biomedical research establishment, especially within the Federal laboratories, is necessary for society to be able to quickly respond to public health crises caused by the appearance of new disease. Recognizing that we support biomedical research both for long-term as well as short-term results, forces attention to a public policy issue that I believe now requires resolution. This is the issue of formulating a decisionmaking framework for managing biomedical research in response to public health emergencies.

To meet the needs of novel health emergencies, like those of the current AIDS epidemic, we must turn to the biomedical research community. There is only one biomedical research community, and it is supported through the policies that this committee is now examining. Thus our "real time" ability to respond to an emergency is inextricably entwined with the policies for support of biomedical research in noncrisis times.

A novel health emergency can be defined as one which attends the appearance of a new disease that is characterized only by unknowns. AIDS is such a novel health emergency. Epidemic polio in the 1930's was a novel health emergency. In both examples, basic biomedical research was first required to define the medical problem as well as the scope of the public health problem.

Research is therefore a mandatory first step in allowing public health officials to develop a national strategy for a novel health crisis. In managing the research response, I believe that the intramural programs of the National Institutes of Health and the Centers for Disease Control have a critical responsibility and play a unique role.

Analysis of the first 6 years of the AIDS experience allows research management problems to be identified that will be shared by other novel health emergencies in the future. The key issues concern the coordination of basic research and the establishment of a national health emergency agenda. As the emergency evolves, there must be flexibility to continually redirect research efforts based on a coordinated and interdisciplinary analysis of the evolving experimental data. To accomplish this requires centralization of management, oversight, and the ability to rapidly shift research directions. These characteristics are more compatible with the structure of the national laboratories than with the diffuse extramural programs.

Although approximately 11 percent of the NIH research budget is expended on intramural programs, these are the programs that are the initial line of response during a novel health emergency. Effective response will always require redirection of ongoing research programs. Initially, providing resources for research is strictly a matter of reallocating funds and talented manpower. The primary policy decisions require assigning responsibility for scientific oversight and instituting program change. The adverse impacts of resource reallocation during emergencies will be minimal if the responsible institutions like the Centers for Disease Control and the National Institutes of Health are solidly funded when the crisis develops.

As any reallocation of resources is disruptive, the goal should be to draw from a broad pool of scientific expertise while minimizing disruption. Initial efforts should thus occur within intramural programs with the advice of extramural experts. The most effective mechanisms for providing technical oversight of research and for capitalizing on expert advice during public health emergencies have yet to be developed. Throughout the emergency, there is need to maintain centralized oversight of the research progress. There is also a need to recognize that various basic research activities merge with all patient care and public health activities. Thus the definition of research during a health emergency is very broad.

Development of public policy in my opinion is needed in four areas to improve the ability of the research community to respond most effectively to a novel health crisis. First, greater budgetary flexibility and discretion within the offices of the NIH and CDC directors will improve the response. Second, a decisionmaking framework for use in the emergency should be developed. An emergency management plan should provide for clear assignment of leadership responsibility, budget authority, centralized management, and coordination. Third, mechanisms for more effectively gathering and synthesizing the expertise of the extramural scientific community are needed. Fourth, new mechanisms for funding extramural crisis research need to be developed.

At present we do not yet have a general process for organizing the Federal research response to a novel health crisis. Each crisis is managed on an ad hoc basis. This is a public policy issue which I believe deserves priority on the national biomedical policy agenda. Recently, Secretary Bowen has been quoted as saying that the management of public health crises is an issue needing attention. I believe that the resolution of this question will force examination

of how we support research as well as acknowledge the important services that the national research laboratories provide. These questions are closely linked to the broader concerns of the role of government in biomedical research that this committee has been examining.

I greatly appreciate the opportunity to express these views.

[The prepared statement of Dr. Panem follows:]

Statement to the Taskforce on Science Policy

of the Committee on Science and Technology

U. S. House of Representatives

by

Sandra Panem*

The Brookings Institution

Wednesday, May 14, 1986

*The views expressed in this statement are mine alone and should not be ascribed to the trustees or staff members of the Brookings Institution.

I am Sandra Panem, a Guest Scholar at the Brookings Institution. I am pleased to be able to appear before this committee and to participate in your discussions concerning the role of the federal government in conducting and supporting basic and applied biomedical research. I commend the committee for convening these hearings. They can only increase appreciation of the importance of federal support for biomedical research, as well as the need for accountability by the scientific community in response to that support.

I have been asked to comment on issues affecting the capability of of the life sciences research community to function in an optimum fashion. In anticipation of this hearing, I have examined the testimony of many distinguished members of the biomedical science community who appeared on April 22, 23 and 24, 1986. Two central issues correctly emerge from that diverse testimony. The first is whether current levels of federal support for biomedical research are sufficient to take maximum advantage of the potential of contemporary science. The second is whether current policies allow the most efficient expenditure of resources for biomedical research -- regardless of the level of funding.

I need only briefly comment on these issues as I am in general agreement with the majority of the other witnesses. The technological advances of the genetic revolution -- now more than a decade old -- have provided biomedical research opportunities that have not yet begun to be tapped. The limiting factors for progress in research are adequate numbers of trained personnel, sufficient laboratory facilities and predictable and continuous funding. Thus the limiting factors are not scientific, but are issues of social choice. Such questions are

eminently suited to resolution in the congress. There can be no question that biomedical science can profitably utilize increasing resources in the foreseeable future for training, for maintaining capital facilitites and for research.

The second central issue is how to foster efficent use of any level of funding. Here, new public policy is needed to foster efficiency by providing consistent and predicable funding. This will allow better planning and eliminate the inefficiencies of disrupting ungoing research before results can be gathered, as well as the inefficiency of time lost in stopping and then starting up research programs. Implementing policies which would allow a multiyear budget cycle for the NIH and CDC would help acheive the goal of more efficient use of resources.

As I read the committee report which lays out the mandate for this congressional Science Policy Taskforce, several underlying assumptions clearly emerge. There is a belief that investment in basic biomedical research is justified because improvements in health derive from research findings. As basic research is therefore in society's interest, it is appropriate for the federal government to support basic research in government facilitites as well as in universities, research institutes, and other organizations. These ideas have guided federal support of biomedical research since the 1930s. Thus, we have traditionally supported research as an investment for the future. I endorse each of these ideas. Yet there is an additional and powerful reason to vigilantly support the basic research enterprise that has not received sufficient emphasis. A strong basic biomedical research establishment -- especially within federal laboratories -- is necessary

for society to quickly respond to public health crises caused by the appearance of new diseases. Recognizing that we support basic research for both long-term as well as short-term results, forces attention to a public policy issue that now requires resolution. This is the issue of formulating a decision making framework for managing biomedical research in response to public health emergencies.

To meet the needs of novel health emergencies -- like those of the current AIDS epidemic -- we must turn to the basic research community. There is only one basic research community, and it is supported through the policies that the committee is now examining. Thus the "real time" ability to respond to such emergencies is inextricably entwined with the policies for support of biomedical research in non-crisis times.

A novel health emergency can be defined as one which attends the appearance of a new disease that is characterized only by unknowns. AIDS is a novel health emergency. Epidemic polio in the 1930s was a novel health emergency. In these examples, basic biomedical research was first required to define both the medical problem as well as the scope of the public health problem.

Research is a mandatory first step in allowing public health officials to develop a national strategy for a novel health crisis. In managing the crisis research response, I believe that the intramural research programs of the NIH and CDC have a critical responsibility and play a unique role.

Analysis of the AIDS experience allows research management problems which will be shared by other novel health emergencies to be identified. The key issues concern the coordination of basic research and the establishment of a national health emergency agenda. As the

emergency evolves, there must be flexibility to continually redirect research efforts based on a coordinated and interdisciplianry analysis of the evolving experimental data. To accomplish this requires centralization of management, oversight and the ability to rapidly shift research directions. These characteristics are more compatible with the structure of national laboratories than with the diffuse extramural programs.

Although only approximately 11% of the NIH research budget is expended on intramural programs, these are the programs that are the initial line of response during a novel health emergency. Effective response will always require re-direction of ungoing research programs. Initially, providing resources for research is strictly a matter of reallocating funds and talented manpower. The primary policy decisions require assigning responsibility for scientific oversight and instituting program changes. The adverse impacts of resource reallocation during emergencies will be minimal if the responsible institutions -- like the CDC and the NIH -- are solidly funded when the crisis develops.

As any reallocation of resources is disruptive, the goal should be to draw from a broad pool of scientific expertise while minimizing disruption. Initial efforts should thus occur within intramural programs with the advice of extramural experts. The most effective mechanisms for providing technical oversight of research and for capitalizing on expert advice during public health emergencies have yet to be developed. Throughout the emergency, there is a need to maintain centralized oversight of the research process. There is also a need to recognize that basic research activities merge with all patient care

and public health activites. Thus the definition of biomedical research during a health emergency is very broad.

Development of public policy is needed in four areas to improve the ability of the research community to respond most effectively to a novel health crisis. First, greater budgetary flexibility and discretion in the offices of the NIH and CDC directors will improve the response. Second, a decision making framework, for use in the emergency should be developed. An emergency management plan should provide for clear assignment of leadership responsibility, budget authority, centralized management and coordination. Third, mechanisms for more effectively gathering and synthesizing the expertise of the extramural scientific community are needed. Fourth, new mechanisms for funding extramural crisis research should be developed.

At present we do not have a general process for organizing the federal research response to a novel health crisis -- each crisis is managed on an ad hoc basis. This is a public policy issue which I believe deserves priority on the national biomedical policy agenda. Secretary Bowen has been quoted as saying that the managment of public health crises is an issue needing attention. I believe that the resolution of this question will force examination of how we support research as well as acknowledge the important services the national research laboratories provide. These questions are closely linked to the broader concerns of the role of government in biomedical research that the committee has been examining.

I greatly appreciate the opportunity to express these views.

DISCUSSION

Mr. FUQUA. Thank you very much.

Discussing our ability to respond to emerging public health crises, you emphasized a need for a strong intramural research program. How do you determine a balance between the intramural and the extramural research? Is the NIH about right? Or the Centers for Disease Control? Is there a strong intramural program and maybe a coordination between the two?

Dr. PANEM. I think that these are difficult questions, and I believe that there is perhaps not a single answer which would allow us to write an equation for the correct distribution between resources intramurally and extramurally. I made my comments because I believe that recent debate for the past 10 years has focused almost exclusively on the number of awards and the level of funding within the extramural communities. It is not to minimize the importance of that community from which we have all greatly benefited, but rather to indicate that there is a need in discussing the balance to recognize that the centralized capability of national research laboratories allows an organization of response which we cannot expect in a diffuse community. Therefore, I believe that the debate concerning the distribution of resources is more properly viewed as a total debate, and not fractured into a discussion between the extramural versus the intramural programs.

Clearly the intramural programs, which represent approximately 11 percent, can never possibly duplicate all of the effort that is located within the extramural community. Therefore, I think it is incumbent upon us to have a better integration of the direction of those intramural programs by allowing the development of new mechanisms to synthesize the advice of the extramural community. I think that that can be improved, and that is worthy of discussion.

Mr. FUQUA. There are agencies like the Centers for Disease Control that have the responsibility to respond to medical emergencies. Of course we had the Legionnaires, AIDS, and it seems like maybe they did respond to an emergency. Many times those are very difficult to predict. I remember back in the mid-1970's, I think, when we were anticipating, I think it was the flu epidemic. That was anticipated and we made all the vaccine, and it never developed. And here again people were trying to be prepared.

Now is there any way you can anticipate what an emergency is going to be?

Dr. PANEM. I think if we were able to anticipate it wouldn't be an emergency.

Mr. FUQUA. Right.

Dr. PANEM. So I think we start from the assumption that we need to be able to be flexible in order to respond to something which is developing. I think your comments concerning the swine flu incident of the 1970's is also of interest, not the least because when Legionnaires disease developed it was in fact thought to be the onslaught, the first onslaught of the swine flu epidemic which did not materialize.

However, I think that those examples are cogent because it recognizes that when one has to respond to something which is unknown one can only reallocate from ongoing programs. And I be-

lieve that the responsibility of the Federal laboratories are different from the extramural programs because there is a responsibility to minimize disruption which I think can be best done when we can have centralized oversight, which is more appropriate within the Federal laboratories.

Mr. FUQUA. Is there any area of biomedical research policy that you think needs particular improvement, particularly in future years?

Dr. PANEM. I think that some of the questions which you will be touching in part later this morning speak to what I think is the most difficult problem that requires real expanded discussion throughout the community, as well as among decisionmakers who will allocate resources, and that is the question of what constitutes appropriate scientific oversight. I believe that this is important because the opportunities which are available should the scientific community have the resources to apply them are far greater than we could possibly accommodate regardless of the increasing resources that one might argue would become available. So consequently, I believe that these decisions become one of policy choices which require improved priority setting which relates to scientific oversight, how we collect the information, how we synthesize it, and how we develop forums in which one can develop a consensus among both the scientific community and within the society as to where we want to apply these opportunities.

Mr. FUQUA. What is the appropriate congressional oversight, rather than congressional interference?

Dr. PANEM. I believe—and this is clearly a very personal opinion. I believe that congressional oversight is appropriate over process. I think that there is a distinct difference between the process by which one comes to decisions rather than the actual selection of scientific programs which I believe would be better established by debate within the scientific community, broad debate, and I do not necessarily believe that the mechanisms which we currently have available are ones which allow full discussion of those questions.

Mr. FUQUA. What about issues such as, a most recent one, AIDS? The Congress determined through a lot of education and public awareness and concern that that had the potential to be a very serious problem, particularly since we didn't know that much about it, or the public did not. And Congress put more money last year into AIDS research. That was not requested by—it was later requested by the administration, but only after Congress had started putting it into the budget.

Now would that be appropriate response? We didn't say where it should go or how it should be spent, but it was for a broad—well, I guess it is not broad, but an area of real national concern.

Dr. PANEM. I would first want to commend the Congress for taking the lead in providing the resources in the recent AIDS epidemic. I think the fact that that was the way in which those resources were generated speaks to the fact that we have not yet developed a vehicle whereby we can agree: first, that circumstances allow an emergency to be declared; and second, that we have established a process by which we will determine where the resources shall come. So those are points where I think we could improve

both the congressional mechanisms which are existent as well as formalize the intent to be able to respond.

The question of what constitutes scientific oversight of the items on the research agenda versus what looks like allocation for resources in specific line items is one which I believe we would not reach closure on here. However, I do believe that it is one that requires discussion. And I think that especially during an emergency the kinds of adversarial discussions which are frequently held within the Congress are not necessarily the best use of time. Consequently, I think that an examination of process and a development of a mechanism where perhaps one could eliminate some of the checks and balances which work so well in noncrisis times could be eliminated.

Mr. FUQUA. Thank you.

Mr. Lewis.

Mr. LEWIS. Thank you, Mr. Chairman. Mr. Chairman, you hit on some of the high points that I wanted to discuss with Dr. Panem.

I am glad that you used the AIDS problems that we have as a novel health emergency, and I would like to just discuss with you a little bit how we could develop a policy in the four areas that you point out in your paper on page 5. I happen to believe that too often Congress gets involved. I think that we did take a gigantic step last year by coming up with the additional funding for the AIDS program.

Having a small city in my district in Florida which has captured all the headlines throughout the country in most of the major magazines about the AIDS epidemic has really caused some very serious social problems, as well as medical problems, in the area.

How do we establish an emergency management plan, in your opinion, to attack this particular problem? We have a number of priorities at the NIH, and several other organizations have as to what they consider their priorities, but in my judgment, many of my colleagues, many of the physicians that I talk to, this could be the most critical emergency that we have that we are sleeping on top of and it is going to come up and grab us before we do something about it. We don't know enough about it. We get into political arguments within the health emergency areas ourselves, as well as the health units themselves, as to what is a priority to pursue.

How would you feel, or what do you think is a plan, or an emergency management plan that we could expedite to attack this AIDS problem?

Dr. PANEM. Thank you for those questions. I would initiate my comments by saying that clearly every health emergency is unique, and consequently, any discussion or formulation of an emergency management process can only set up a framework which will ease the management of idiosyncratic decisions which will be related to each emergency, as well as to, I believe, force attention to some of the social problems which perhaps do not come naturally and under basic biomedical research or public health mandates. In my opinion, the most difficult problem with developing such a framework recognizes first that the institutions which we must turn to, the public health institutions, the research institutions, the social service institutions, are organized independent of each other, as are

the oversight committees and the authorization and appropriations processes for the Federal agencies, and many of these services are best organized at the State and local levels. Thus I think what an emergency management plan would do would be to establish the precedence and the routes by which many of the turf and organizational barriers which allow many of the issues which need to be integrated during a health emergency and which are normally separated to be crossed.

Second, I believe that one could then assign a clear individual who would have responsibility and in whose mandate would be the requirement to look both at the social, the fiscal, the research, and the social service problems. I believe that if one brings to the table the people who are representatives of the public administration community, the political leadership, the scientific and the social service community we could design what would be a health emergency equivalent of some of the emergency management programs which have been established within FEMA. I think that that is something which could be done.

I think the more difficult question, which is the one that I believe requires initiation of discussion now, is how to establish the criteria which would allow us to utilize that system. And I believe that the current AIDS epidemic is one which indicates that, had we a process, some of the problems of whose authority and whose responsibility that we have experienced in the past 6 years could be perhaps eased. And it is the reason that I argue that a committee such as this which is fundamentally concerned with the infrastructure to which we must turn can also help by asking how do we organize discussions within this community under extreme conditions as well as under noncrisis times.

I am sorry that I am not more specific. I think it is a question which really requires broad social debate.

Mr. LEWIS. Do you feel that we need an organization or task force type or something set up in our National Institutes of Health or somewhere in our infrastructure that would establish a criteria to attack these type of emergencies and specifically pinpoint their efforts in those areas when you have a crisis research, like I feel we should have in AIDS, and I don't think we have it at this point in time? How do you get yourself to that point where you can have direct application into crisis research and continue it until you start making some headway rather than getting into some of the things that you mentioned—the budget battles, or we are going to cut back on this, or as soon as you make your first giant step into the research, well then you have funding cut, where you get into again the turf battles as to whose responsibility it is and the normal scientific disagreements as to opinion as to whether you are heading in the direction or not?

Dr. PANEM. I do think that we need a process that is established. I think that it is one which should probably be located within the Department of Health and Human Services, most probably under the direct purview of the Secretary, most likely to be led by the Assistant Secretary for Health because of the tradition of that individual being a health professional.

I think that how we establish that process and the trigger by which it becomes employed will require debate. I think that debate

can be initiated in a number of ways. One might be through a congressional task force. One might be by the appointment of an executive task force. One might be by commissioning a study, for example, in an institution such as the Institute of Medicine. I think there are a number of ways to initiate the discussion.

I think the result of the discussion might well be an expanded version of the Health Emergency Act with a better definition of the triggers as to how that emergency fund would be tapped, as it has not yet been.

Mr. LEWIS. Thank you, Dr. Panem.

Thank you, Mr. Chairman.

Mr. FUQUA. Thank you very much. We appreciate you being here this morning.

Our next panel of witnesses will be Mr. Walter W. Stewart, a research scientist at the National Institutes of Health, and Dr. Ned Feder, a research scientist at the National Institutes of Health.

Gentlemen, we are delighted to have you with us this morning. And, Mr. Stewart, we would be pleased to hear from you first.

[A biographical sketch of Mr. Stewart and Dr. Feder follow:]

Name: Walter W. Stewart.

Date and place of birth: March 29, 1945. New York, N.Y.

Positions held: 1963-1967 Student, Harvard college; 1967-1968 Graduate Fellow, Rockefeller University; 1968-1972 Commissioned Office, National Institutes of Health; 1972-1974 Junior Fellow, Society of Fellows, Harvard University, Guest Scientist, National Institutes of Health; 1974-1979 Staff Fellow, National Institutes of Health; 1979-1981 Research Physicist, GS-12, National Institutes of Health; and 1981-present Research Physicist, GS-13, National Institutes of Health.

Honors: Consulting Editor, *Nature* (1971-1973); Phi Beta Kappa (Senior year); Summa Cum Laude B.A. (Harvard); and Junior Fellow, Society of Fellows.

Patents: U.S. patent number 4,473,693 (Sept. 25, 1984). Aminonaphthalimide dyes for intracellular labelling.

Awards: Bausch and Lomb Science Medal (Harvard); Decatur Prize (Harvard); Harvard Chemistry Department Award; and Inventor's Incentive Award, U.S. Department of Commerce, 1980.

Invited lectures: West Coast Nerve Net, Honorary lecturer, 1979; Marine Biological Laboratories, 1984; and NATO colloquium, Banyuls sur Mer, course instructor, 1984.

Name: Dr. Ned Feder.

Date and place of birth: April 14, 1928. Minneapolis, Minnesota.

Education: 1948 A.B. (cum laude), Harvard College; 1949 M.S. (chemistry), Harvard Graduate School; and 1953 M.D., Harvard Medical School.

Appointments: 1949-50 Teaching Fellow in Chemistry, Harvard University; 1950 Research Assistant, Department of Nutrition, Thorndike Memorial Laboratory, Boston City Hospital; and Department of Biology, Harvard University (summer); 1951 Research Asssistant, Department of Biology, Harvard University (summer); 1953-54 Intern, Philadelphia General Hospital; 1953-55 Fellow in Biology (American Cancer Society), Department of Biology, Harvard Univesity; 1955-58 Senior Assistant Surgeon, U.S.P.H.S., National Institute of Allergy and Infectious Diseases, National Institutes of Health; 1958-1961 Special Trainee of the National Institute of Neurological Diseases and Blindness, U.S. Public Health Service, at Harvard Medical School; 1961-66 Assistant Professor of Biology, Harvard University; 1966-67 Lecturer, Harvard University; and 1967-pres. Chief, Section on Biophysical Histology, NIADDK, NIH.

STATEMENTS OF WALTER W. STEWART, RESEARCH SCIENTIST, NATIONAL INSTITUTE OF DIABETES AND DIGESTIVE AND KIDNEY DISEASES, NATIONAL INSTITUTES OF HEALTH, BETHESDA, MD; AND DR. NED FEDER, RESEARCH SCIENTIST, NATIONAL INSTITUTE OF DIABETES AND DIGESTIVE AND KIDNEY DISEASES, NATIONAL INSTITUTES OF HEALTH, BETHESDA, MD

Mr. STEWART. Thank you very much, Mr. Chairman. My name is Walter Stewart. We welcome the opportunity to appear before the Science Policy Task Force of the House Committee on Science and Technology. We are pleased and proud to testify before a group interested in "maintaining America's leadership in science in view of the changing environment facing us over the coming decades."

We shall describe here our unpublished study on the frequency of misconduct in a group of 47 biomedical scientists. It is usual to make the assumption that carelessness, irresponsibility, and dishonesty are very rare in the conduct of science. But in the group we studied, misconduct was common. We do not know whether misconduct will prove to be widespread or exceptional when other groups are studied.

Out study provides the first data that bear on certain question. How many scientists engage in misconduct that substantially weakens or even invalidates the reported results of their research? Is the scientific community taking adequate measures to deal with the problem of misconduct? If there is a problem, what can be done to fix it? These are not abstract questions that deal simply with scientific methodology.

We want to emphasize, Mr. Chairman, that both in our unpublished report and in our testimony here, we are presenting our own views, and not acting as spokesmen for the National Institutes of Health. We shall limit our comments to the issues covered in our study.

The two of us, Ned Feder, my colleague, and myself, are scientists who work as full-time Federal employees in a laboratory at the National Institutes of Health. We synthesize chemical compounds that are used to study the shape of nerve cells, and we carry out basic research on the genetic control of nerve cell shape, using isogenic snails as our experimental animal. We decided to carry out the study noted above because we thought we had a novel idea about a little-known subject of great importance.

There are virtually no data on the frequency of misconduct among scientists. In the absence of such data, for example, a science administrator stated in an article, "Like most people in the scientific community, we at the NIH have regarded the incidence of real or apparent scientific misconduct as of low frequency * * *." The opposite view, however, is also widely held. For example, a biologist has suggested that "unconscious or dimly perceived finagling, doctoring, and massaging are rampant, endemic, and unavoidable * * *." But there have been virtually no supporting facts for either point of view.

It is striking to us that there are virtually no data on a subject obviously of great interest and social importance. A leading journal of medical research recently published the proceedings of a symposium in which the five participants discussed various forms of sci-

entific misconduct. The editor stated bluntly in an accompanying editorial that "the frequency of abuses and waste is not known."

In the spring of 1983, Ned Feder and myself thought of a way to measure the frequency of professional misconduct in a large sample of scientists. The idea was to study a particular group of 47 biomedical scientists about whose professional practices an unusual amount of published information was available. The members of the group had been coauthors of a scientist who, unknown to the people in our sample, had forged much of the data on which their publications were based. The forgery was investigated by two committees appointed by the deans at Emory and Harvard Medical Schools, where the research had been done, and by a committee appointed by the National Institutes of Health, the Federal agency that had paid for most of the research at Harvard.

The public reports of the committees that investigated the forgery contained much information about the coauthors: their private research practices were described in the reports, often in detail. It was for this reason that we chose to study this particular sample. I might mention that the problem, as you may have gathered all along, is that it is not easy to get empirical data on this subject, and the idea we had was a particular way whereby we could get empirical data on a group.

We described the idea that we had to the editor of *Nature*, a leading scientific journal, and he expressed interest in publishing such a study. In the next 5 months we obtained the documents we wanted to study, analyzed them, wrote a report on our research, and submitted it to *Nature*. We would like to submit for the record a draft copy of our report, version of May 11, 1986, as a part of our testimony. Our report is still unpublished, for reasons we will describe later. My colleague Ned Feder will now describe something about the sample and the methodology that we used.

Dr. FEDER. The coauthors whom we studied ranged in experience from novices to department chairmen who had been involved in research for many years. None were known to us, either personally or professionally; their field of research was completely different from our own. Working with the scientist who had forged data, they had coauthored 109 scientific publications over a period of 3 to 4 years. We did not name any of the coauthors in our report, but a reader could deduce their identities by studying the published documents on which our report was based.

The publications by those in our sample were in the fields of clinical and experimental cardiology. A typical clinical paper might contain a description of a particular form of heart disease in a group of patients or a description of the diagnosis or treatment of a particular heart disease. The typical experimental paper might describe the testing and comparison of new drugs administered to dogs soon after an experimentally induced heart attack.

We studied the professional practices of the 47 authors as reflected in the reports of the investigating committees and also in the coauthors' publications in scientific journals. We examined only the publications they had coauthored with the scientist who had forged data, not their other publications. We were especially careful to distinguish between the original episode of forgery, which

was not the subject of our report, and the practices directly attributable to the coauthors themselves.

In the report that we wrote on our research, we listed in tabular form all the examples of scientific misconduct we could find in our sample. Altogether there were 10 categories. We also gave detailed descriptions of many specific examples of scientific misconduct. Any reader could go to the nearest science library and check the accuracy of most of these descriptions; the rest could be checked by examination of the committee reports.

One common practice we discussed in our report was the publication of errors and discrepancies in papers coauthored by those in our sample. We found these errors, an average of 12 per research paper, simply by reading the papers carefully; a knowledge of cardiology was not needed. We held the coauthors responsible for failing to check the papers that bore their own names. We considered this to be what we called a type A practice; one explainable simply by carelessness or excessive haste. Some examples of careless errors appeared spectacular, for example, a published family tree from which one could calculate the ages at which the men and women had become parents. It is easy to show that one of the males had supposedly fathered two children by the time he was 9 years old. Many careless errors were quite minor, for example, a small discrepancy between a number in a table and what should have been the same number in the text. The important point is not simply that there were errors, but that there were so many of them.

Another type A practice was what we called honorary authorship. This occurs when a scientist, usually the head of a laboratory or clinic, puts his name on a paper when he is not directly involved in the research. Our evidence for this practice came from the reports of the investigating committees. We think that honorary authorships falsify the assignment of responsibility for published research and increase the likelihood that inaccurate data will be published.

There was also scientific misconduct that was more serious, in our opinion, and that we called type B. In early versions of our report, we characterized these as unethical practices. In more recent versions, we described them instead as involving an apparent lack of candor. For example, some coauthors published statements that we believe, at the time of publication, they knew or should have known were inaccurate. As we discuss in our report in detail, the conditions of the experiment, including the method of choosing the control animals, were described misleadingly and inaccurately in the published papers in such a way as to make the data and the conclusions look better to the reader than they actually were. Much of our evidence for this practice came from the report of one of the investigating committees. The coauthors acknowledged at the time of the investigation that, as laboratory policy, some controls had been selected in a fashion different from that described on publication, and there is additional, similar evidence that they knew or should have known of the inaccurate description of the experimental conditions.

In our report we described and tabulated a variety of other type A and type B practices, for example, failure to retain patients'

names in such a way that access to clinical data would be permitted, unacknowledged republication of data, and so on.

Here is a summary of our findings. We found that considerably more than half the scientists in our sample, 35 out of 47, had engaged in questionable practices, according to our analysis. We also concluded that one-fourth of our sample, 13 out of 47, had engaged at least once in the more serious type of misconduct, type B.

The reported results of the scientists' published research, their findings, were in some instances weakened or even invalidated by the effects of their misconduct.

It is important to emphasize that the sample of scientists we studied was not chosen in a random fashion and thus our findings may or may not be representative of those that would be obtained in other groups of scientists or in the scientific community at large. Other studies will be needed to determine this. However, our study does raise the possibility that misconduct may be far more common among biomedical scientists than is realized.

In writing our report, we considered it desirable to include enough information so that our analysis could be verified. From the literature references that we provided for this purpose, it is possible for a strongly motivated reader to go to the library, do some work and deduce the identity of those in our sample. But the singling out of individual scientists was not our intention. The more important questions, in our view, are, first, the high frequency of misconduct in the entire group, and second, the implications of this finding for science in general.

And here is what happened when we tried to publish our report. Two scientists in our sample hired lawyers and threatened a libel suit against us and against any journal that published our report. The editors of major scientific journals, *Nature* and *Cell,* stated that they considered our report important and worthy of publication. However, these journals failed to publish our report after considering it for 16 months and 9 months, respectively.

More recently, for the past 8 months, *Nature* has again been considering our report for publication. We have been told by that journal that our report will be published soon. But in view of the previous difficulties with publication, we cannot assert with confidence that this will occur.

We also sent our report to 14 other scientific journals for informal consideration by the editors. Although more than half the editors commented favorably on our report, there was not a single journal for which publication appeared at all likely. Accordingly, we decided not to make a formal submission to any of the latter journals.

It seems to us likely that fear of a libel suit was an important consideration, but not the sole one, in the editors' decision not to publish our report. Some editors may have felt that our report, whose conclusions and implications are quite unsettling, might not be well received by some of their readers. It is worth noting that the editor of *Science,* a prominent international journal published by the American Association for the Advancement of Science, told us, during a telephone conversation in which he indicated that our report was unsuitable for publication in his journal, that *Science*

would be giving less coverage to scientific misconduct in the future than it has in the past.

I would like to turn now to some possible causes of misconduct.

Misconduct among scientists may perhaps be related in some way to their previous behavior in school. In most studies, roughly 30 to 60 percent of students are found to cheat in high school, college, graduate school, and medical school. It seems unlikely that those who successfully cheat as students are somehow excluded from a career in research, or that they enter research but turn over a new leaf.

Another possible cause may be the extraordinary emphasis commonly placed on sheer quantity of publications, regardless of accuracy, quality, or probable usefulness. The director of one of the world's leading research institutions sent a signed, official memorandum proposing incentives for publication to some scientists working in his organization. Parts of the memorandum read as follows:

> There is no demand that these be literary masterpieces in first line journals; journeyman works for publication in second, third or fourth line archival publications will be quite satisfactory. Upon proper completion and submission of [two of these] manuscripts, [the appointment of a technician now working for you] will be extended to April 1, 1984. During that time it is expected that an additional manuscript on [a particular subject] will be completed and submitted. If so, the period of employment will be extended an additional three months and again an additional manuscript on [the subject] is an anticipated result of the extended employment.

On another occasion a director of research said in all seriousness, "No research that results in a single paper is worthwhile."

There are also several indications in our own study that the pressure to publish may have led to some of the questionable practices that we found.

The extreme competition for research grants and the phenomenon of team science may also contribute to the genesis of scientific misconduct. In our opinion, some who are skillful at acquiring authority may value and reward conformity over adherence to sound research practices. One writer has used the term "hired hand research." When the team leader's expectations differ from the facts, the junior scientist on the team faces a hard choice.

I would like to comment next on a subject we consider to be quite important. The essence of honorary authorship, as we defined it earlier, is the claiming of scientific credit where virtually no scientific contribution has been made. Often this occurs when a senior scientist, who controls the funds for a research group that he directs, obtains credit by authorship for scientific contributions that are in fact minimal. This misattribution of scientific credit and the widespread tolerance of the practice have a number of harmful effects that go beyond the unfairness of taking credit where it is not due.

First, the combination of power and remoteness of the chief of the laboratory may lead him to make a poor choice of research projects, to insist that they be carried to completion, perhaps with inappropriate methods, and finally to make an incorrect judgment of the accuracy of the research results. Nevertheless, his high position lends a prestige that may lull or intimidate other coauthors,

the reviewers, or the readers into uncritical and inappropriate acceptance of the work as valid.

Second, sheer numbers of publications accumulated by a laboratory chief become a measure of his success as a scientist, and he is likely to apply the same criterion in judging other scientists in his own laboratory and elsewhere, with the harmful consequences I have mentioned previously.

Third, the illusion is created that the chief is doing research, when in reality he is not. The so-called principal investigator, as the holder of research grants is formally designated, may in these cases not really be an investigator at all. We emphasize that these comments apply solely to research that is done by a team of scientists led by a scientist-administrator not actually involved in the research.

It is probably no accident that many of the discovered instances of fraud have occurred in large research groups. Such groups offer a setting conducive to scientific misconduct.

The absence of a written code of conduct for scientific research may also facilitate scientific misconduct. The editor of a leading journal of medical research told us after reading a draft copy of our report that we failed to make allowance for what he called the "subtleties and ambiguities of real life." No one, he said, lives by what we call the accepted standards of research.

Mr. STEWART. Mr. Chairman, I would like to make a comment. Just listening to this, it sounds so grim in certain ways, that I would just mention what you said in the beginning: that we were studying not science as a whole but what was wrong in the sample we found. We are just describing the things we found in our sample, and because the sample is not representative, one just doesn't know how well these findings are going to generalize.

There are various ways in which self-regulation is supposed to operate in science. All the coauthors of a research report are generally considered to take responsibility for part or all of the research itself as well as for the published report describing that research. Furthermore, the editor of a scientific journal customarily asks scientists to serve as reviewers of the submitted reports and advise him on their accuracy and importance. Finally, after research is published, other scientists may check the work by attempting to repeat it.

How well does the system of scientific self-regulation really work? No one knows, actually. In some cases of outright forgery of data, the customary methods of checking published research have worked well, and the problem was recognized quickly and it was corrected quickly. It is especially likely, we believe, that forgeries of important results will be rapidly detected, simply because there will be attempts at repetition right away. But there is no reason to believe that scientific self-regulation consistently detects or discourages the less flagrant and presumably more frequent forms of questionable practices in science.

As one distinguished scientist commenting on our study noted, "Our present arrangements in science have an unfortunate tendency to reward those who publish voluminously, and the correcting influence of replication and criticism simply cannot keep up with the magnitude and complexity of the resulting mass of publica-

tions. Thus the danger of detection of carelessness, sloppiness, or even worse is relatively slight. The time lag required for [attempts at replication] is often years or even decades, while decisions about grants, jobs, promotions, and so forth, must be made over a much shorter time span."

Although the coauthors in our sample may or may not be statistically representative of biomedical scientists in general, this is less likely for the journal referees and others who were responsible for checking the work. These papers were published in leading biomedical journals despite their many defects, and this suggests that the usual mechanisms of peer review thought to protect the accuracy of the scientific literature may be less effective than is commonly believed.

In the episode of forgery discussed in our study, the response of the scientific community proved inadequate even though there was clear evidence of serious misconduct. Investigating committees were appointed at both medical schools, but the performance of these committees suggests that this mechanism for studying alleged misconduct—a committee chosen by the dean largely from within his own institution—may not be effective.

At one of the medical schools, the investigating committee concluded in its report that none of the forged data had been published, but the facts proved otherwise. A committee appointed by NIH later found that seven of the eight research papers published from that medical school by the scientist in question were, in fact, contaminated by forgery. The investigating committee at the other medical school concluded, after obtaining statements from the coauthors of a particular published paper, that the paper, which contains recommendations for drug therapy of patients, was valid. We examined the paper and without great difficulty found that it contained so many errors that it is almost certainly invalid.

Given the importance of the subject, it is unfortunate that there are almost no published studies on the frequency of misconduct among scientists. And we are referring here to actual empirical data—where people have actually tried to measure it. In our opinion this may be because there is almost no institutional support for such studies. Based on the comments of other scientists, as well as our own personal experience, we think that if such studies are begun attempts may be made to interrupt or censor them. If completed, they may be unacceptable to the editors or publishers of scientific journals. As a consequence, even those scientists who regard the subject as interesting and important may avoid working on it.

It appears to us probable that most scientists have not given much thought to the practical or ethical implications of the practices described here. One colleague of ours described these as "little scientific white lies." And it is possible that the success of many scientists may depend on their continuing to engage in these practices.

It seems possible that the same pressures that lead susceptible scientists to engage in misconduct may have pervasive, harmful effects on the way that almost all scientists plan and carry out their work. In a conversation with one of us, a scientist at one of the Nation's leading medical schools sheepishly referred to much of his own work as "semimeaningful research," which he explained as re-

search largely lacking in importance but chosen because it could be counted on to produce publishable results. And it is our impression that this practice is common.

Harm is done by scientific misconduct. The end product of scientific research is information communicated to other scientists or used in some application. If that information is degraded, whether through carelessness or deceit, the effect may be a waste of resources and a slowing of scientific progress, both in basic and applied research.

There may be a direct connection between professional misconduct by biomedical scientists and delays in discoveries leading to the prevention or treatment of disease. What is unknown is the extent of the misconduct and the magnitude of the resulting delays.

In the course of carrying out our study, we have had many private discussions with senior scientists. A large fraction of them have encountered instances of serious professional misconduct in their own laboratories, but have taken little or no action for fear of the legal or social consequences. This suggests to us that serious professional misconduct may be more common than is currently realized, and further suggests that institutional mechanisms for dealing with this problem are inadequate.

We have distinguished between type A practices, which are explainable simply by carelessness or haste, and type B practices, which involve an apparent lack of candor. Many readers of our report have argued that, in the harm done to science, outright fraud is in some sense the worst. They add that type B practices are more urgently in need of attention than those of type A. But it seems possible to us that type A practices are more common than those of type B and that outright fraud is the least common. This suggests a surprising conclusion to us. In their aggregate effect, type A practices, that is, those that are less serious, may in fact be more harmful than those of type B. It is even possible that outright fraud, in its aggregate effect, may do the least harm of all to science.

We discuss here—since time is running short I will abbreviate a little bit—some previous figures given on the incidence of misconduct in science, but these were based on self-reporting. Our figures differ quite markedly from those figures, and no one knows at present what the actual figures are for sure. However, the figures that come from our study are clearly not tolerable, and if they are in fact widespread they pose a profound threat to the integrity of scientific research and the scientific literature.

Responses to our report. We found the responses of our colleagues quite revealing. We received criticisms and suggestions from more than a hundred scientists. Aside from three particular people whose practices were by implication discussed in our report and the lawyers representing two of these scientists, virtually no one said that our description of facts was inaccurate. And I should qualify that now as I see it. We did receive notice of a certain number of inaccuracies, maybe 10 or 12, as we passed it out to these 100 people and, of course, we fixed those. And that was one good reason to show it to so many people. Most readers made favorable comments on our study. Many respected senior scientists

stated that they considered it an important service for science and its publication as being in the public interest.

We have given in appendix B excerpts from some 45 readers, both pro and con. And if you have the time or the patience, you can read through them all. I should mention that we tried to put in all of the opinions that we got that were negative, so that we were not balancing it toward favorable opinion.

Dr. FEDER. Not all that were negative, but most of those that were negative.

Mr. STEWART. Thank you.

One point that emerged clearly is that many people found the report very disturbing—many scientists. Surprisingly, most people were not surprised by our findings. That surprised us.

What should be done? Here we have some suggestions, and we should say right away that we have no special competence, none at all, to suggest solutions for this problem. We think our data point convincingly to the existence of a problem, and we have no idea really how to solve it.

The problems we have discussed are serious and obviously require attention. The most pressing problem is to remedy the deficiencies we have described.

It would probably also be worthwhile to get a more precise numerical measure of the extent of misconduct in science. This could be done in various ways. One is an audit of a random sample of published scientific papers in order to determine what fraction of these papers are demonstrably defective at the time of publication. And in particular, we had in mind papers which are published but for which there is no data existing at the time of publication to support these papers. An audit of this sort might either include examination of the raw data and the laboratory notebooks, or it could be limited simply to the careful review of published research reports, which was, in essence, the methodology we used. The first type of audit where you go in and look at the raw data and the laboratory notebooks is more comprehensive. The advantage of the second is that it is less intrusive.

Obviously, auditing of this sort has its costs, of which the financial costs might be the least. Unwarranted harm to individual scientists and a general atmosphere of suspicion created by auditing could be costly in the long run to the functioning of science.

The practice of profiting from research one has not actually done appears to be centrally involved in the origin of many of the practices we observed in this study. Accordingly, it seems reasonable to consider ways of reasserting the principle that scientific credit be given only for scientific work actually accomplished. No single approach is likely to solve this complex problem. Some journals are now requiring from authors a statement that defines their participation in submitted papers. In our report we suggest another possible approach: decreased scientific support for large research groups directed by a single chief, and correspondingly increased support for individual scientists. This would not prevent the individual scientists from joining together in teams wherever that seemed appropriate for them and advantageous for the investigation of particular scientific problems.

Mr. Chairman, it seems unnecessary to wait for further studies to measure the level of scientific misconduct before taking steps to deal with it. Our data point to a problem that is substantial, but whose exact size is not known. However, we feel that it is not too soon for scientists to start considering solutions to the problem, regardless of its exact extent. Our findings clearly indicate the need for scientists, especially the leaders of science, to pay more attention to the validity of published research and the accuracy of the scientific literature.

We are hopeful that our colleagues will, by solving these complex problems, justify the continuing confidence of the public in the scientific establishment and maintain America's preeminence in the biomedical sciences.

Mr. Chairman, this concludes our statement. We would be pleased to answer any questions you or anyone else may have.

[The prepared statement of Mr. Stewart and Dr. Feder follows:]

For release upon delivery

Statement of Walter W. Stewart
and Ned Feder, M.D.
National Institute of
Diabetes and Digestive and Kidney Diseases
National Institutes of Health
Department of Health and Human Services

Before the
Task Force on Science Policy
of the Committee on Science and Technology
U. S. House of Representatives

May 14, 1986

63-471 - 87 - 2

Mr. Chairman, we welcome the opportunity to appear before the Science Policy Task Force of the House Committee on Science and Technology. We are pleased and proud that we were asked to testify before a group interested in "maintaining America's leadership in science in view of the changing environment facing us over the coming decades."

We shall describe here our unpublished study on the frequency of misconduct in a group of 47 biomedical scientists. It is usual to make the assumption that carelessness, irresponsibility, and dishonesty are very rare in the conduct of science. But in the group we studied, misconduct was common. We do not know whether misconduct will prove to be widespread or exceptional when other groups are studied.

Our study provides the first data that bear on certain questions: how many scientists engage in misconduct that substantially weakens or even invalidates the reported results of their research? Is the scientific community taking adequate measures to deal with the problem of misconduct by scientists? If there is a problem, what can be done to fix it? These are not abstract questions that deal simply with scientific methodology.

We want to emphasize, Mr. Chairman, that both in our unpublished report and in our testimony here, we are presenting our own views, and not acting as spokesmen for NIH. We shall limit our comments to the issues covered in our study.

The two of us are scientists who work as full-time federal employees in a laboratory at the National Institutes of Health. We synthesize chemical compounds that are used to study the shape of nerve cells, and we carry out basic research on the genetic control of nerve cell shape, using isogenic snails as our experimental animal. We decided to carry out the study noted

above because we thought we had a novel idea about a little-known subject of great importance.

There are virtually no data on the frequency of misconduct among scientists. In the absence of such data, for example, a science administrator stated[1] in an article, "Like most people in the scientific community, we at the NIH have regarded the incidence of real or apparent scientific misconduct as of low frequency . . . ." The opposite view is also widely held. For example, a biologist has suggested[2] that "unconscious or dimly perceived finagling, doctoring, and massaging are rampant, endemic, and unavoidable . . . ."* But there have been virtually no supporting facts for either position.

It is striking that there are virtually no data on a subject obviously of great interest and social importance. A leading journal of medical research recently published the proceedings of a symposium in which the five participants discussed various forms of scientific misconduct. The editor stated bluntly in an accompanying editorial[3] that "the frequency of abuses and waste is not known."

In the spring of 1983 we thought of a way to measure the frequency of professional misconduct in a large sample of scientists. The idea was to study a particular group of 47 biomedical scientists about whose professional practices an unusual amount of published information was available. The members of the group had been coauthors of a scientist who, unknown to them, had forged much of the data on which their publications were based. The forgery was investigated by two committees appointed by the deans at Emory and Harvard medical schools, where the research had been done, and by a committee appointed by NIH, the federal agency that had paid for most of the research at Harvard. The public reports of the committees that investigated the forgery contained much information about the coauthors: their private research practices were described in the reports, often in detail. It was for this reason that we chose to study this particular sample.

---

*References for these quotations and other citations are given in Appendix A.

We described the idea to the editor of Nature, a leading scientific journal, and he expressed interest in publishing such a study. In the next five months we obtained the documents we wanted to study, analyzed them, wrote a report on our research, and submitted it to Nature. We would like to submit for the record a draft copy of our report (version of 5-11-86) as a part of our testimony. Our report is still unpublished, for reasons we shall explain later.

The sample that we studied

The coauthors ranged in experience from novices to department chairmen who had been involved in research for many years. None were known to us, either personally or professionally; their field of research was completely different from our own. Working with the scientist who had forged data, they had coauthored 109 scientific publications. We did not name any of the coauthors in our report, but a reader could deduce their identities by studying the published documents on which our report was based.

The publications by those in our sample were in the fields of clinical and experimental cardiology. A typical clinical paper might contain a description of a particular form of heart disease in a group of patients or a description of the diagnosis or treatment of a particular heart disease. A typical experimental paper might describe the testing and comparison of new drugs administered to dogs soon after an experimentally induced heart attack.

Subject of the report

We studied the professional practices of the 47 coauthors as reflected in the reports of the investigating committees and in the coauthors' publications in scientific journals. We examined only tne publications they had coauthored with the scientist who had forged data, not their other publications. We were especially careful to distinguish between the original episode of forgery, which was not the subject of our report, and the practices directly attributable to the coauthors.

In the report we wrote on our research, we listed in tabular form all the examples of scientific misconduct we could find in our sample. Altogether there were ten categories. We also gave detailed descriptions of many specific examples of scientific misconduct. Any reader could go to the nearest science library and check the accuracy of most of these descriptions; the rest could be checked by examination of the committee reports.

One common practice we discussed in our report was the publication of errors and discrepancies in papers coauthored by those in our sample. We found these errors — an average of twelve per research paper — simply by reading the papers carefully; a knowledge of cardiology was not needed. We held the coauthors responsible for failing to check the papers that bore their own names. We considered this to be what we called a type A practice: one explainable simply by carelessness or excessive haste. Some examples of careless errors appeared spectacular, for example, a published family tree from which one can calculate the ages at which the men and women had become parents; it is easy to show that one of the males had supposedly fathered two children by the time he was nine years old. Many careless errors were quite minor, for example, a small discrepancy between a number in a table and what should have been the same number in the text. The important point here is not simply that there were errors, but that there were so many of them.

Another type A practice was what we called "honorary authorship." This occurs when a scientist — usually the head of the laboratory or clinic — puts his name on a paper when he is not directly involved in the research. Our evidence for this practice came from the reports of the investigating committees. We think honorary authorships falsify the assignment of responsibility for published research and increase the likelihood that inaccurate data will be published.

There was also scientific misconduct that was more serious, in our opinion, and that we called type B. In early versions of our report, we characterized these as "unethical practices." In more recent versions, we described them instead as involving an "apparent lack of candor." Some

coauthors published statements that we believe, at the time of publication, they knew or should have known were inaccurate. As we discuss in our report in detail, the conditions of the experiment, including the method of choosing the control animals, were described misleadingly and inaccurately in the published papers, in such a way as to make the data and the conclusions look better to the reader than they actually were. Much of our evidence for this practice came from the report of one of the investigating committees. The coauthors acknowledged at the time of the investigation that, as laboratory policy, some controls had been selected in a fashion different from that described on publication, and there is additional, similar evidence that they knew or should have known of the inaccurate description of experimental conditions.

In our report we described and tabulated a variety of other type A and type B practices: failure to retain patients' names that would permit access to clinical data, unacknowledged republication of data, and so on.

Our conclusions

We found that considerably more than half the scientists in our sample (35 out of 47) had engaged in questionable practices, according to our analysis. We also concluded that about one-fourth of our sample (13 out of 47) had engaged at least once in the more serious type of misconduct (type B).

The reported results of the scientists' published research were in some instances weakened or invalidated by the effects of their misconduct.

It is important to emphasize that the sample of scientists we studied was not chosen in a random fashion and thus our findings may not be representative of those that would be obtained in other groups of scientists or in the scientific community at large. Other studies will be needed to determine this. However, our study does raise the possibility that misconduct may be far more common among biomedical scientists than is realized.

In writing our report, we considered it desirable to include enough information so that our analysis could be verified. From the literature references provided for this purpose, it is possible for a strongly motivated reader to deduce the identity of those in our sample. But the singling out of individual scientists was not our intention. The more important questions, in our view, are the high frequency of misconduct in the entire group, and the implications of this finding for science in general.

Unsuccessful attempts to publish our report

Two scientists in our sample hired lawyers and threatened a libel suit against us and against any journal that published our report. The editors of major scientific journals, Nature and Cell, stated that they considered our report important and worthy of publication. However, these journals failed to publish our report after considering it for 16 months and 9 months, respectively. More recently, for the past 8 months, Nature has again been considering our report for publication. We have been told by that journal that our report will be published soon. In view of the previous difficulties with publication, we cannot assert with confidence that this will occur.

We also sent our report to 14 other scientific journals for informal consideration by the editor. Although more than half the editors commented favorably on our report, there was not a single journal for which publication appeared at all likely. Accordingly, we decided not to make a formal submission to any of the latter journals.

It seems to us likely that fear of a libel suit was an important consideration, but not the sole one, in the editors' decision not to publish our report. Some editors may have felt that our report, whose conclusions and implications are quite unsettling, might not be well received by some of their readers. It is worth noting that the editor of Science (a prominent international journal published by the American Association for the Advancement of Science) told us, during a telephone conversation in which he indicated our report was unsuitable for publication in his journal, that Science would be giving less coverage to scientific misconduct in the future than it has in the past.

A more detailed description of our unsuccessful attempts at publication was presented by us on February 26, 1986, in testimony before the House Judiciary Subcommittee on Civil and Constitutional Rights.

In our report we discussed several issues related to our findings: causes of misconduct, self-regulation in science, the harm done by scientific misconduct, and possible ways of dealing with the problem.

Possible causes

Although the causes of misconduct among scientists are not known, some may be suggested.

Misconduct among scientists may perhaps be related in some way to their previous behavior in school. In most studies[4-10] roughly 30-60% of students are found to cheat in high school, college, graduate school, and medical school. It seems unlikely that those who successfully cheat as students are somehow excluded from a career in research, or that they enter research but turn over a new leaf.

Another possible cause may be the extraordinary emphasis commonly placed on sheer quantity of publications, regardless of accuracy, quality, or probable usefulness. The director of one of the world's leading research institutions sent a signed, official memorandum proposing incentives for publication to some scientists working in his organization: "There is no demand that these be literary masterpieces in first line journals; journeyman works for publication in second, third or fourth line archival publications will be quite satisfactory. Upon proper completion and submission of [two] manuscripts, [the appointment of a technician now working for you] will be extended to April 1, 1984. During that time it is expected that an additional manuscript on [subject] will be completed and submitted. If so, the period of employment will be extended an additional three months and again an additional manuscript on [subject] is an anticipated result of the extended employment." On another occasion a director of research said in all seriousness, "No research that results in a single paper is worthwhile."

There are several indications in our study that the pressure to publish may have led to many of the questionable practices we found. Coauthors who published more papers tended to have more of certain types of scientific misconduct per paper than coauthors who published less ($r=0.51$; $0.01<p<0.05$). Also, in letters of recommendation written about the scientist later found to have forged data, there was generally more emphasis on the quantity of his publications than on their quality, and 3 of 4 such letters referred prominently and approvingly to his curriculum vitae, which lists the publications and their titles, but contains no information about the accuracy or importance of their contents.

The extreme competition for research grants and the phenomenon of team science may also contribute to the genesis of scientific misconduct. In our opinion, some who are skillful at acquiring authority may value and reward conformity over adherence to sound research practices. One writer[11] has used the term "hired hand research." When the team leader's expectations differ from the facts, the junior scientist on the team faces a hard choice.

The essence of honorary authorship, as we defined it previously, is the claiming of scientific credit where virtually no scientific contribution has been made. Often this occurs when a senior scientist, who controls the funds for a research group that he directs, obtains credit by authorship for scientific contributions that are in fact minimal. This misattribution of scientific credit and the widespread tolerance of the practice have a number of harmful effects that go beyond the unfairness of taking scientific credit where it is not due.

— First, the combination of power and remoteness in the chief of the laboratory may lead him to make a poor choice of research projects, to insist that they be carried to completion, perhaps with inappropriate methods, and finally to make an incorrect judgment of the accuracy of the research results. Nevertheless, his high position lends a prestige that may lull or intimidate other coauthors, the reviewers, or the readers into uncritical and inappropriate acceptance of the work as valid.

— Second, sheer numbers of publications accumulated by a laboratory chief become a measure of his success as a scientist, and he is likely to apply the same criterion in judging other scientists in his own laboratory and elsewhere, with the harmful consequences we have discussed previously.

— Third, the illusion is created that the chief is doing research, when in reality he is not. The so-called Principal Investigator (as the holder of research grants is formally designated) may in these cases not really be an investigator at all. (We emphasize that these comments apply solely to research that is done by a team of scientists led by a scientist-administrator not actually involved in the research.)

It is probably no accident that many of the discovered instances of fraud have occurred in large research groups. Such groups offer a setting conducive to scientific misconduct.

The absence of a written code of conduct for scientific research may facilitate scientific misconduct. The editor of a leading journal of medical research told us after reading a draft copy of our report that we failed to make allowance for the "subtleties and ambiguities of real life"; no one, he said, lives by what we call the accepted standards of research.

Self-regulation: scientists watching scientists

There are various ways in which self-regulation is supposed to operate in science. All the coauthors of a research report are generally considered to take responsibility for part or all of the research itself as well as for the published report describing that research. Furthermore, the editor of a scientific journal customarily asks scientists to serve as reviewers of the submitted reports and advise him on their accuracy and importance. Finally, after research is published, other scientists may check the work by attempting to repeat it.

How well does the system of scientific self-regulation really work? No one knows. In some cases of outright forgery of data, the customary methods of checking published research have worked well, and the problem was recognized and corrected quickly. It is especially likely that forgeries of important results will rapidly be detected. But there is no reason to believe that scientific self-regulation consistently detects or discourages the less flagrant and presumably more frequent forms of questionable practices in

science. As one distinguished scientist commenting on our study noted, "Our present arrangements in science have an unfortunate tendency to reward those who publish voluminously, and the correcting influence of replication and criticism simply cannot keep up with the magnitude and complexity of the resulting mass of publications. Thus the danger of detection of carelessness, sloppiness, or even worse is relatively slight. . . . [T]he time lag required for [attempts at replication] is often years or even decades, while decisions about grants, jobs, promotions, etc. must be made over a much shorter time span. . . . "

Although the coauthors in our sample may not be statistically representative of biomedical scientists in general, this is less likely for the journal referees and others who were responsible for checking the work. These papers were published in leading biomedical journals despite their many defects, and this suggests that the usual mechanisms of peer review thought to protect the accuracy of the scientific literature may be less effective than is commonly believed.

In the episode of forgery discussed in our study, the response of the scientific community proved inadequate even though there was clear evidence of serious professional misconduct. Investigating committees were appointed at both medical schools, but the performance of these committees suggests that this mechanism for studying alleged misconduct — a committee chosen by the dean largely from within his own institution — may not be effective. At one of the medical schools, the investigating committee concluded in its report that none of the forged data had been published; a committee appointed by NIH later found that 7 of the 8 research papers published from that medical school by the scientist in question were contaminated by forgery. The investigating committee at the other medical school concluded, after obtaining statements from the coauthors of a particular published paper, that the paper (which contains recommendations for drug therapy of patients) was valid; we examined that paper and without great difficulty found it to contain so many errors that it is almost certainly invalid.

Given the importance of the subject, it is unfortunate that there are almost no published studies on the frequency of misconduct among scientists. In our opinion this may be because there is almost no institutional support for such studies. Based on the comments of other scientists[15], as well as our own personal experience, we think that if such studies are begun, attempts may be made to interrupt or censor them; if completed, they may be unacceptable to the editors or publishers of scientific journals. As a consequence, even those scientists who regard the subject as interesting and important may avoid working on it.

It appears to us that most scientists have not given much thought to the practical or ethical implications of the practices described here; one scientist referred to these practices as "little scientific white lies." Indeed the professional success, if not survival, of many scientists may depend on their continuing to engage in these practices, which often earn them substantial rewards. Many see these practices as privately sanctioned by those in authority. For example, one writer[11] has referred to workers who "modify or fabricate portions of the reports in order to give the boss what he seems to want."

It seems possible that the same pressures that lead susceptible scientists to engage in misconduct may have pervasive, harmful effects on the way that almost all scientists plan and carry out their work. In a conversation with one of us, a scientist at one of the nation's leading medical schools sheepishly referred to much of his own work as "semi-meaningful research," which he explained as research largely lacking in importance, but chosen because it could be counted on to produce publishable results. It is our impression that this practice is common.

Harm done by scientific misconduct

The end product of scientific research is information communicated to other scientists or used in some application. If that information is degraded — whether through carelessness or deceit — the effect may be a waste of resources and a slowing of scientific progress, both in basic and applied research.

There may be a direct connection between professional misconduct by biomedical scientists and delays in discoveries leading to the prevention or treatment of disease. What is unknown is the extent of the misconduct and the magnitude of the resulting delays.

In the course of carrying out our study, we had many private discussions with senior scientists. A large fraction of them have encountered instances of serious professional misconduct in their own laboratories, but have taken little or no action for fear of the legal or social consequences. This suggests to us that serious professional misconduct may be more common than is currently realized, and further suggests that the institutional mechanisms for dealing with this problem are inadequate.

We have distinguished between type A practices, which are explainable simply by carelessness or excessive haste (for example, failure to correct errors), and type B practices, which involve an apparent lack of candor (for example, deliberate misstatements). Many readers of our report have argued that, in the harm done to science, outright fraud is the "worst," in some sense. They add that type B practices are more urgently in need of attention than those of type A. But to us it seems reasonable that type A practices are more common than those of type B, and that outright fraud is the least common. This suggests a surprising conclusion to us: in their aggregate effect, type A practices may be more harmful than those of type B. It is even possible that outright fraud, in its aggregate effect, may do the least harm of all to science.

The frequency of misconduct among biomedical scientists is still an open question. Three science administrators[12-14] have previously given figures indicating a frequency of detected scientific misconduct of about 0.1%. In our sample of 47 scientists we found a frequency of almost 70%. Even if only the type B practices are considered, the figure is almost 30% — still several hundredfold higher than the previous estimate. It is clear that the earlier figures and our own cannot simultaneously represent the frequency of scientific misconduct. The figures on frequency of misconduct that emerge from our study are clearly not tolerable and, if widespread, pose a profound threat to the integrity of scientific research and the scientific literature.

Responses to our report

The responses to our report were quite revealing. We received criticisms and suggestions from more than a hundred scientists. Aside from three scientists whose practices were discussed in our report and the lawyers representing two of these scientists, virtually no one said that our description of specific facts was inaccurate. Most readers made favorable comments on our study. Many respected senior scientists stated that they considered it an important service for science and its publication as being in the public interest.

Excerpts from the comments of 45 readers are given in Appendix B.

One point emerged more clearly from telephone conversations with readers than from the written comments: our report appeared to be disturbing to almost all scientists. Most were not surprised by its findings and regarded it as accurate.

What should be done?

The problems we have discussed are serious and obviously require attention from the scientific community. The most pressing problem is to remedy the deficiencies we have described.

It would probably also be worthwhile to determine the extent of misconduct in science. This could be done in various ways. One is an audit of a random sample of published scientific papers in order to determine what fraction of these papers are demonstrably defective at the time of publication. An audit might either include examination of raw data and laboratory notebooks, or could be limited simply to the careful review of published research reports. The first type of audit is more comprehensive; the advantage of the second is that it is less intrusive.

Auditing has its costs, of which the financial costs might be the least. Unwarranted harm to individual scientists and a general atmosphere of

suspicion created by auditing could be costly in the long run to the functioning of science.

The practice of profiting from research that one has not actually done appears to be centrally involved in the origin of the many practices we observed in this study. Accordingly, it seems reasonable to consider ways of reasserting the principle that scientific credit be given only for scientific work actually accomplished. No single approach is likely to solve this complex problem. Some journals are now requiring from authors a statement that defines their participation in submitted papers. In our report we suggest another possible approach: decreased scientific support for large research groups directed by a single chief, and correspondingly increased support for individual scientists. This would not prevent the individual scientists from joining together in teams wherever that seemed appropriate for them and advantageous for the investigation of particular scientific problems.

Mr. Chairman, it seems unnecessary to wait for further studies to measure the level of scientific misconduct before taking steps to deal with it. Our data point to a problem that is substantial, but whose exact size is unknown. However, we feel it is not too soon for scientists to start considering solutions to the problem, regardless of its exact extent. Our findings clearly indicate the need for scientists, especially the leaders of science, to pay more attention to the validity of published research and the accuracy of the scientific literature.

We are hopeful that our colleagues will, by solving these complex problems, justify the continuing confidence of the public in the scientific establishment and maintain American preeminence in the biomedical sciences.

Mr. Chairman, this concludes our statement. We would be pleased to answer any questions you may have.

Appendix A: References

Frequency of misconduct:

1. Raub, W. F. (1981). The National Institutes of Health's perspective on misconduct by grantees. In: Whistleblowing in Biomedical Research, J. P. Swazey and S. R. Scher, eds. (Washington: U. S. Government Printing Office), pp. 151-154.

2. Gould, S. J. (1978). Morton's ranking of races by cranial capacity. Science 200, 503-509.

3. Huth, E. J. (1986). Abuses and uses of authorship. Annals of Internal Medicine 104, 266-267.

Cheating in school:

4. Bowers, W. J. (1964). Student Dishonesty and Its Control in College (New York: Bureau of Applied Social Research, Columbia University).

5. Harp, J. and Taietz, P. (1966). Academic integrity and social structure: A study of cheating among college students. Social Problems 13, 365-373.

6. Zastrow, C. H. (1970). Cheating among college graduate students. Journal of Educational Research 64, 157-160.

7. Sierles, F., Hendrickx, I., and Circle, S. (1980). Cheating in medical school. Journal of Medical Education 55, 124-125.

8. Baird, J. S., Jr. (1980). Current trends in college cheating. Psychology in the Schools 17, 515-522.

9. Schab, F. (1980). Cheating among college and non-college bound pupils, 1969-1979. Clearing House 53, 379-380.

10. Pemberton, C. (29 November 1983). Institutional research study 83-32: Results from the spring 1983 student and faculty surveys on academic honesty at the University of Delaware, pp. 1-32.

Hired hand research:

11. Roth, J. A. (1966). Hired hand research. American Sociologist 1, 190-196.

Frequency of misconduct is 0.1%:

12. Malone, T. E. (19 April 1983). "Fostering Integrity in Research: A Collective Responsibility." Presentation in Dallas, Texas, pp. 1-10.

13. Brandt, E. N., Jr. (1983). PHS perspectives on misconduct in science. Public Health Reports 98, 136-140.

14. Miers, M. L (1985). Current NIH perspectives on misconduct in science. American Psychologist 40, 831-835.

Others:

15. Personal communications.

Appendix B: Some comments on our report

We quote here from letters or written statements on our report. The great majority of comments we received were favorable. However, we have included in this appendix, for the committee's information, a disproportionately large number of the unfavorable comments.

Comments by editors

— "I think that what you have done is an important public service . . . ." (Editor of Nature. October 21, 1983.) "Let me explain that from my point of view the overriding need is not so much to modify your paper so as to avoid a libel suit but to ensure that it is both fair and the scientific document it purports to be." (Editor of Nature. January 11, 1985.)

— "The consensus of opinion has been that it would be overwhelmingly in the public interest for this article to be published." (Editor of Cell. September 4, 1985.)

In addition to Nature and Cell, we sent our report to 14 other scientific journals for informal consideration by the editor. All rejected it or made comments indicating that acceptance for publication was quite unlikely. Excerpts from letters of six of these editors follow.

— "The Associate Editors felt, in general, that some of the conduct you illustrate in the paper is widely known in the academic community, which is not to say that it is conduct that is regarded as admirable. So, in essence, our decision comes down to a matter of relative priorities. . . . Your study was thorough and admirable." (March 15, 1985.)

— "I do not concur with you that the details of this case are extraordinarily important or that they will affect public policy on fraud.

Obviously, from what you report, many other editors feel the same way I do." (July 23, 1985.)

— "The basis for your sample was co-authoring with Darsee. This leaves it completely indeterminate whether or not this population is representative of the larger biomedical research community and its practices. Notice that one can draw neither conclusion: neither that the practices are typical and represent a general problem, nor that they are specific to Darsee collaborators and represent a specific case only." (August 8, 1985.)

— "The legal ramifications will obviously cause any editor to pause. But I am sure you will have achieved one thing. You have probably made a lot of editors think deeply about the professional and ethical standards of scientists and about how to write on them. I assure you that you have done that to me." (August 26, 1985.)

— "I found the paper to be overstated, wooly, and excessively judgemental and moralistic. . . . On the other hand, the paper does contain some disturbing and in my opinion unassailable points." (A letter from one journal editor to another. August 30, 1985.)

— ". . . I am sorry to write that we see no feasible way to use it. . . . Both our publisher and our long-time legal advisor believe that the legal costs incurred would be prohibitive. I am disappointed that we have to send you this negative decision, and that there is nothing we can do to help. It has been sobering to look at your and Dr. Feder's clear and convincing report." (September 6, 1985.)

In addition to the possibility of a libel suit, there were other reasons given by the editors for their unwillingness to publish our report.

— The report was too long (but none offered to print it if it were shortened).

— The subject of the report was not suitable for that journal.

— Professional misconduct was already well known to be common among scientists; it was unnecessary to write about it, and probably damaging to the scientific community to do so.

— We had exaggerated the seriousness of misconduct by some of the coauthors in our sample, and we had made the general problem of misconduct appear worse than it actually is.

None said that our description of specific facts was inaccurate.

Comments by anonymous reviewers chosen by NIH

The NIH had our report reviewed by three anonymous reviewers. Excerpts from the comments of all three are included below.

— "I believe that this is an important paper and I would recommend publication. I think that 'the conclusions flow from the data,' and the results are 'new and interesting.' . . . . If I had to make the decision myself, I would probably publish it with the 'imprimatur of NIH' but again with an accompanying letter from the Director of NIH and the Director of NIADDK explaining why we preferred to have the paper prepared by the authors 'on their own time.'" (April 2, 1984.)

— "(1) I think this manuscript contains the substance for a worthwhile paper. The topic is important, the data are interesting, and the thrust of the conclusions follows directly from the findings. However, the phrasing of the conclusions needs work (see below). (2) I believe the manuscript fails in one significant respect; it is a hybrid of technical writing and science journalism and, as such, is a good example of neither. For example, the data are presented with unnecessary editorialization when, in fact, they speak rather well unvarnished . . . . (3) If the authors agree to revise the manuscript to eliminate the deficiencies indicated above, I would favor allowing them to complete it and submit it to a major journal as part of their normal duties. It seems appropriate for intramural scientists to concern themselves about the practices within their profession so long as they do not interfere with or allow themselves to be mistakenly associated with NIH investigations into possible misconduct." (April 2, 1984.)

— "This paper is in fact two papers. The first deals with a very specific example of misconduct on the part of one investigator, and shows how his collaborators and associates were drawn into lending their names (at least) to this misconduct. This first paper is factual (except where it confidently attributes to the collaborators motives that can only be guessed at). . . . The second paper contained in the work of Stewart and Feder is quite different. It is a speculation, based on the incident they have analyzed, to the effect that dishonesty among biomedical scientists may be widespread, that it is a condition to be expected on the basis of the known behavior of medical students, and that the correct way to deal with this problem is to undertake a more general investigation of the honesty of all such scientists. I think there is no scientific basis in the preceding part of the manuscript for such a speculation; it is entirely gratuitous." (June 18, 1984.)

Comments by anonymous reviewers chosen by a journal editor

— "I feel very strongly that it would be in the public interest to publish this manuscript, with or without some slight modifications, and I know that its contents would be interesting to scientists and to others concerned with research." (Anonymous reviewer #1.)

— "I have reviewed page by page of the manuscript by Stewart and Feder. This is obviously a highly charged and extremely sensitive subject. The material painful as it is to face has been presented very well and I think fairly. The major thrust of this article it seems is to illustrate the many ways in which investigators may tend to be entrapped wittingly or unwittingly into carrying out fraudulent or scientifically unsound research. The issue of accountability is particularly well stressed. The authors were careful to omit virtually every single name in the Darsee affair other than Darsee and the committee heads or members responsible for institutional review after the discovery of fraud. This attempt to protect many individuals who may or may not be innocent is laudable. . . . Despite its length, I believe this manuscript must be published in order that the medical and bioscientific professions give evidence of their willingness to face the fact that fraud may

be committed and that there are reasons as well as motivations how this may occur." (Anonymous reviewer #2.)

— "The authors do not discuss in sufficient detail what the papers they analyze are examples of. . . . The second problem with the paper is that the authors set their own standards for what is and is not acceptable scientific practice. . . . The second part of the paper which deals with setting standards and protection against fraud also makes me uneasy. I do not disagree with much of what Stewart and Feder propose; indeed how could one? I do find the notion of a study involving audits, internal and external, chilling but could understand that others would find it acceptable." (Anonymous reviewer #3.)

— "Should this paper be published? . . . Published anywhere? My answer, yes. I believe that some of the journals who published the original papers, especially several of them, should publish this analysis. Their apparent unwillingness to do so is scandalous and reflects poorly on their editors, and the scientific specialty they represent; to some extent it reflects poorly on all scientists." (Anonymous reviewer #4.)

— "I have no doubt that it would be in the scientific and public interest for Stewart and Feder's analysis to be published. . . . Stewart and Feder's paper is obviously in an inappropriate format for Cell. That is a disappointment since its subject concerns invalid publishing practices. Like Caesar's wife, this paper must itself be above suspicion, and the presentation of the tables which are difficult to read and interpret does not aid external auditing. . . . On the other hand, one more frequently comes across scientists whose work one simply 'does not trust.' I have had one such person in my laboratory with whom I refused to publish or allow a paper to be published under the name of the institution where we then worked. But to accuse him a fraud without firm evidence would have been libelous. So he is still 'at large' causing similar difficulties in a series of U.S. laboratories, though not as a result of any recommendations I have made orally or in writing." (Anonymous reviewer #5.)

— "I believe that this is an important paper that should certainly be published somewhere. . . . It seems clear that Stewart and Feder are determined that this material will see the light of day. If they are unable to obtain its publication in any reputable scientific journal, this will be a further indictment of the scientific community." (Anonymous reviewer #6.)

— "I do not believe that the paper is appropriate for Cell because the selected subset consists of individuals in the fields of medicine, psychology and animal physiology — all very remote from the readership of Cell. With only a few exceptions, the cited work was published in journals that have zero overlap with the Cell readership, and which as far as I can tell may have entirely different standards of quality and stringency of review. . . . The paper is written in a tedious style which will not appeal to the busy readership of Cell. I would reserve space for more important things." (Anonymous reviewer #7.)

— "In the long run, the public interest will benefit from publication of this manuscript. . . . The potential negative influence of this article on the careers and reputations of Darsee's co-authors, as well as Stewart and Feder, is such that were this a trivial topic I would strongly discourage its publication. The topic of this article is not trivial. There are many who would prefer to ignore what we would like to think is a very rare occurrence of unprofessional conduct among medical investigators. I take the opposite view. The topic of acceptable 'Professional practices among biomedical scientists' should be discussed repeatedly and openly. . . . In summary, this is a clearly written report of research in an area that is important to the scientific community. Where facts are claimed, they appear to be adequately documented such that the reader can personally verify the claims." (Anonymous reviewer #8.)

Comments by deans or presidents of major universities and medical schools

— "I do not feel that the work of Stewart and Feder falls into the category of a constructive contribution, although I found no major

misrepresentation or mistakes in the manuscript. . . . In conclusion, I am uncertain what useful purpose this exercise by Stewart and Feder accomplishes and I do [not] think it a manuscript of the quality usually found in Nature." (Interim Dean of a leading medical school to Editor of Nature. April 4, 1984.)

"Quite honestly, I think this manuscript is a tragedy and I would strongly urge that it not be published. It strikes me as petty, vindictive, gratuitous, self-serving, and of no useful purpose." "I believe that it is a great disappointment that these two men would take it upon themselves to level criticisms at very distinguished scientists [i.e., Drs. Richard Ross, Daniel Tosteson, and Neil Moran] who accomplished the tedious, laborious, and painful examinations of the Darsee papers . . . . Perhaps even more regrettable is the fact that these authors have chosen to editorially assassinate some respected members of the faculty in two medical schools." "For my part, I certainly do not wish to engage in any public exchange with these two individuals; in this country, there is a crude admonition against engaging in competitive micturitional exhibitions with Mephitis mustelidae (American translation available upon request). Thank you for letting me see this scurrilous piece of work. It is distressing at best, abhorrent for the most part. I sincerely hope that no one will publish this vicious diatribe . . . ." (President and Acting Dean of a leading medical school to Editor of Nature. April 9, 1984.)

— ". . . in my opinion, the paper submitted by Stewart and Feder does little or nothing to increase understanding of these issues [i.e., the determinants of dishonesty and fraud]. . . . This paper contains many allegations and interpretations with which I disagree. I do not think that it would be helpful to you for me to comment on these points in detail." (Dean of a leading medical school to Editor of Nature. April 13, 1984.)

— "In short, what you have — or so it seems to me — is an interesting in-depth appendix to an important case of scientific fraud. Your manuscript, however, tries to convert it into something quite different. In my judgment, it fails to do so convincingly." (President of a leading university. February 26, 1986.)

Comments by other scientists

The following comments are taken from letters sent to us by 22 other readers. Most are senior scientists, and most are members of the National Academy of Sciences.

— "In my judgment this is a serious and responsible inquiry. . . . Apart from the ethical issues of scientific responsibility, which are my primary concern, such practices waste large amounts of Government money . . . . [I]t is a responsible endeavor, which is intended to help in maintaining high standards in research, and is therefore intended to serve the interests of the NIH and the scientific community . . . ."

"In my opinion you have performed a real service, by placing the debate on scientific honesty and fraud on a more objective level. . . . Now that you have broken the ice, I trust that other studies will follow; they are obviously needed." (October 21, 1983, and August 5, 1984.)

— "With [your conclusions] I concur entirely. . . . I also feel that one of your conclusions, namely that we ourselves better do the criticism lest others decide to do so for us, is very appropriate. In short, I think the paper is excellent, and I sincerely hope that Nature will publish it." (February 16, 1984.)

— "The topic addressed is one of great importance to scientists and the public alike and, considering that importance, has been little examined. . . . The study described in this manuscript is a serious, constructive effort to examine a troubling series of events. . . . The work merits publication and widespread discussion, particularly among scientists and academicians." (February 20, 1984.)

— "[T]his is the most detailed review I have ever read of a series of scientific papers. . . . I hope the paper will be printed in its entirety in Nature. It should give very serious cause for thought on the part of many individuals." (July 5, 1984.)

— "No one will venture to deny the facts or contradict your main conclusions." (July 17, 1984.)

— "On the whole you have done a remarkable job and deserve much credit for putting matters that are often repressed or unsupported in a light where that is no longer possible." (August 31, 1984.)

— ". . . I wonder whether your article will lead to rational discourse especially as questions about the publication of inaccurate and republished data, practices in the retention of primary data, qualifications for authorship, the role of collaborators, etc. have been mixed up with other weighty matters such as the correct institutional response to allegations of scientific fraud and misconduct." (December 10, 1984.)

"You do not do a good job of putting the results of your study in perspective. Despite your disclaimers, you imply repeatedly that the practices you ran across are common, even quite common. This implication is likely to stir up enormous resentment and indignation by those who have a strong sense of responsibility to truth. Just present what you know. Let it speak for itself . . . ." (October 30, 1985.)

— "Your paper . . . impresses me as a very thoughtful and thorough analysis of strong evidence of very sloppy practices. . . . The whole matter is very important, and regardless of the embarassment(s) that may result, I feel it is important that other scientists be informed about the facts as you present them. I am therefore even more disturbed by the evidence that fear of the expense of defending against threatened libel suits has apparently caused scientific journals to refuse publication of your paper. This verges on a cover-up. . . . I am afraid that you may have dealt with only the tip of a regrettable iceberg. Our present arrangements in science have an unfortunate tendency to reward those who publish voluminously, and the correcting influence of replication and criticism simply cannot keep up with the magnitude and complexity of the resulting mass of publications. Thus the danger of detection of carelessness, sloppiness, or even worse is relatively slight. . . . [T]he time lag required for [attempts at replication] is often

years or even decades, while decisions about grants, jobs, promotions, etc. must be made over a much shorter time span. . . . I feel you have made an important contribution to science. . . ." (November 19, 1985.)

— "I suggest also that you may have constructed too fine a screen, and I make this suggestion with some reluctance. It would probably be a good thing to review all the complicated matters involved with a view to dealing only with those that are, in the words of the law, 'beyond reasonable doubt,' leaving points that are supported by 'the weight of the evidence,' a much less stringent requirement, for later comment. I suppose I am become too pragmatic about some of the points which you, quite rightly, criticize. Overlapping controls, data in text that don't match data in tables (when they should be identical), rushing into print, and even publishing abstracts in several different places with altered titles may not, in the extremity to which we have come in the last thirty or so years, be looked on as anything more than misdemeanors. But contriving data, stealing it from other authors, and fraudulently manipulating it to reach a conclusion that for some reason is desired are felonious, relatively, and I think your initial treatment should focus on those features. I repeat that I make such a suggestion reluctantly because most of the shortcomings you have attacked are less than honorable; and I dislike writing as pragmatist rather than as purist. . . . Your condemnation of honorary authorship is splendid and I hope it will survive. What's more, I think many editors and researchers will applaud — not condemn — your denunciation of the practice. Likewise the insane emphasis on quantity of publication." (November 26, 1985.)

— "I have read your report and took the trouble to look up a couple of the original articles that you cite. I also discussed your paper with two colleagues. It is interesting, revealing, and sobering." (December 26, 1985.)

— "I think is it an important contribution in that it raises important questions about standards in our profession. I am as appalled as you that lawyers should have a veto on what gets published." (December 31, 1985.)

— "I have read the report with much interest. I think it is important that it be published." (January 3, 1986.)

— "So, although I can appreciate your effort — which was clearly with good intentions — I cannot accept your conclusion as being valid. I would guess that most open-minded scientists would feel the way I do, and only those who want to believe that dishonesty is rampant, will embrace your findings." (January 31, 1986.)

— "Finally, I must state my surprise that you chose to take time off from your laboratory research for many months in order to prepare this manuscript. Perhaps it's a puritanical notion that we should use working hours to perform the work for which we are hired and engage in other activities, however professional and meritorious, on our own time such as evenings and weekends." (February 3, 1986.)

— "I am in complete agreement with the major points you have made in your manuscript. I think they should be published. The question is how, in view of the obstacles you have run into. . . . It is often the task of writers, such as yourself, to say what we all know, but bring it into the open and say it well and forcefully." (February 12, 1986.)

— "As your paper now stands, it just looks in utterly exhausting detail at one very sick episode. And then it seems to generalize from there to all of science. . . . The reason for these lawsuit threats is that most of the people involved around the periphery of this affair feel that they have lived a clean scientific life. Knowing quite a few of them, I tend to agree. And they feel that this type of paper is just 'muckraking' which seeks to generalize about them even to the point of categorizing their personality defect as 'Type A or Type B.'" (February 14, 1986.)

— "I read your report and believe that it would greatly benefit scientific research if it were published. . . . Many scientists have never given this matter much, if any thought, and many are a bit loose about rigorous adherence to strict intellectual honesty. It would be good to have the issues

explicitly exposed as you do in your manuscript. . . . I hope you succeed with publication. It will be a valuable service to all of us." (February 19, 1986.)

— "I have studied with great interest yours and Ned Feder's painstaking, yet objective and discrete analysis of what one could call loose fidelity in many professional practices. We can all learn from your sober analyses." (March 1, 1986.)

— "Although I have not consulted any of the original material or articles, I am deeply impressed with what you have accomplished, and I am convinced that your findings need to be brought to the attention of a wide audience of scientists for careful considerations. . . . I congratulate you on this useful and important piece of work . . . ." (March 7, 1986.)

— "While I have absolutely no professional competence in the area treated your fully documented account leaves little doubt that there is an appalling amount of sloppiness, special pleading, deviousness, and even deceit in the set of papers you chose to consider. . . . [A]ll chemists know of papers which are demonstrably false, not a few of which were probably deliberately faked. Very few of these are ever publicly retracted: research workers simply discount them by experience or by word of mouth. The . . . incident [of a Ph. D. that was revoked] mentioned above was an interesting exception since not only was the degree revoked but a retraction was published in the same journal as the false report had appeared in. Subsequently we learned that the malefactor had done much the same thing at a previous school . . . and yet letters of recommendation chose to ignore the matter. It is not only editors who are pusillanimous." (March 17, 1986.)

— "This experience [i.e., an incident involving an institutional response to dishonesty], the Darsee affair, and the other data you cite suggest the following conclusions:

— Dishonesty is endemic in professional practice.

— The pressures of medical training and biomedical practice provide a climate which facilitates dishonesty. As you suggest, some of the most

successful people are the dishonest ones.

— Medical school admissions committees, academic appointment committees, and journal referees make little or no effort to screen for dishonesty.

— When evidence of dishonesty is found, it is actively suppressed by the academic community." (April 29, 1986.)

Disk WS115 5-11-86

PROFESSIONAL PRACTICES AMONG BIOMEDICAL SCIENTISTS: A STUDY OF A SAMPLE GENERATED BY AN UNUSUAL EVENT

Walter W. Stewart and Ned Feder

National Institute of Diabetes
and Digestive and Kidney Diseases
National Institutes of Health
Bethesda, Maryland 20892

We report here a study on the frequency of professional misconduct in a group of 47 biomedical scientists. Our findings suggest that misconduct among scientist may be several hundredfold more common than believed from some previous estimates. This raises the possibility that misconduct among scientists may be wasting hundreds of millions of dollars annually in the U.S. alone and furthermore may be retarding scientific progress, for example, discoveries leading to the prevention or treatment of disease. What is unknown is the extent of the misconduct and the magnitude of the resulting delays. There are indications that these problems may be associated with large research groups directed by scientist-administrators who do almost no research themselves. We suggest that the improper attribution of scientific credit may be largely responsible for the problems discussed here.

We wish to report the results of an unusual investigation. In May, 1981, the colleagues of a respected young scientist were shocked to discover that he was a forger of data. Initially, Dr John Darsee confessed to a single forgery, but it was subsequently found by his colleagues and by three investigating committees that he had fabricated much of the data that formed the basis for his more than 100 publications[1-129] over a period of about three years.

It occurred to us that Darsee's extraordinary behaviour had unwittingly provided a test -- which could hardly have been done in any other way -- of the customary behaviour of other scientists: a test of his coauthors, a test of the referees and others responsible for the accuracy of the scientific literature, and finally a test of the institutional response to overt fraud. We report here our analysis of the circumstances surrounding this unusual event.

Our study provides the first data that bear on questions of great practical importance. How many scientists engage in misconduct that substantially weakens or even invalidates the reported results of their research? Is the scientific community taking adequate measures to deal with the problem of misconduct

by scientists? These are not abstract questions that deal with scientific methodology. The health and welfare of those who pay the bills for science depend on discovering the answers to these questions.

Our sample consists of the 47 scientists who were coauthors of Darsee. Their practices are described in some detail in this report. They worked at two leading medical schools and ranged in experience from novices to department chairmen. As a by-product of the various investigations of the forgery, their practices were described in detail in the reports of the investigating committees. Because this "sample of convenience" (A#227) was generated by an unusual event and was not chosen in a random fashion, another type of study will be needed to determine whether our results are representative of clinical scientists, of biomedical scientists, or of scientists in general, or alternatively are unrepresentative of any of these groups.

We were astonished by what we found when we analyzed our data. Considerably more than half the scientists in our sample (35 out of 47) had engaged in questionable practices, according to our analysis. We also concluded that about one-fourth of our sample (13 out of 47) had engaged at least once in a serious type of misconduct. The reported results of the scientists' published research were in some instances weakened or invalidated by the effects of their misconduct.

It is possible that our study will prove to be of very limited value. Our sample of scientists -- because of their close association with a known forger of data -- may be quite unrepresentative of other groups of scientists, and our study may shed almost no light at all on the frequency of misconduct among scientists as a whole. Nonetheless, our study raises, in a form difficult to dismiss, the possibility that misconduct is far more common among biomedical scientists than is now realized, with possibly profound consequences for the public health and welfare.

At the end of this report we propose a study to gather further information about the accuracy of the scientific literature. But we believe that at present the prevailing atmosphere and the available resources make such a study almost impossible; certainly there has never been any such study. We suggest specific steps that might be effective in lowering the frequency of scientific misconduct.

Background

The circumstances of the Darsee affair have been widely reported (A#130-133,156-162). In 1978 to 1981 inclusive, he was author or coauthor of 18 full-length research papers published in major biomedical journals, and of about 100 abstracts, book chapters, letters, reviews, and short papers in clinical and experimental cardiology[1-129]. In part because he was so prolific, he was highly regarded by his scientific peers

and superiors, first at Emory University School of Medicine and then at Harvard Medical School.

Five months after Darsee's confession, in May 1981, to a single act of data fabrication, it became apparent that there had been more than one such episode. Three committees were appointed to investigate. Harvard appointed a committee headed by R. S. Ross, which issued its report (A#130) in January 1982. The NHLBI (National Heart, Lung, and Blood Institute) appointed a panel headed by H. E. Morgan to investigate the circumstances at Harvard, where most of the research had been funded by the National Institutes of Health. Their report (A#133) was completed in June 1982. Finally, at Emory, an investigating committee headed by N. C. Moran was formed, and issued its report (A#131,132) in March 1983. The committee reports analyzed Darsee's research and incidentally shed light on some of his coauthors' practices in research and publication.

Shortly after the report of the Moran committee was released, and almost two years after the frauds themselves had come to light, we embarked on our study. The primary data we analyzed were the 109 publications coauthored by the 47 scientists in our sample. We carefully examined these documents, looking for published errors, seemingly misleading statements, and other defects. We also used the reports of the three investigating committees. We were especially careful to distinguish between the direct effects of Darsee's forgeries, which are not the subject of this report, and the defects

attributable to his coauthors. We recognize that some scientists might disagree with some of our decisions as to what constitutes a defect attributable to a coauthor.

Darsee's publications were in experimental and clinical cardiology, a field of research quite different from our own. For this reason, we have no comments to make on the pointedness or importance of the research described in the publications or about the design of the experiments. No doubt we may also have overlooked some errors in the publications because of our lack of first-hand knowledge of the field. At the same time, we believe that our position as outsiders had the advantage that none of those who published with Darsee was known to us personally or professionally.

The sample we studied was composed of Darsee's 47 coauthors. Twenty-four worked at Emory University School of Medicine, and 23 at Harvard Medical School. Biographical data were found for 44 of the 47 coauthors. Thirty-nine were identified as M.D.s, one as a Ph.D., and one as an M.D.-Ph.D; three apparently had no advanced degree. At the time of their involvement with Darsee, about half of the coauthors were 30-40 years old and had received their advanced degree 5-15 years previously. Approximately equal numbers were junior and senior to this group. Those junior included technicians, medical students, resident staff at teaching hospitals, and junior faculty members. The senior coauthors included several professors and department chairmen. (Darsee himself was 33 in

1981.)

Abundance of errors

Most of the published errors and discrepancies were presumably introduced by Darsee. We considered all errors, major and minor, to be of interest because they might reflect on the care with which the coauthors had checked the publications that bore their own names, and might reflect indirectly on the care with which they had carried out their share of the research.

All but a few of the 18 research papers[6,24-26,50-53,87-89,90-96] (Table 1) and many of the abstracts contain errors or discrepancies that can be detected simply by examining them carefully; a knowledge of cardiology was not needed. We emphasize again that many of the errors and discrepancies are minor, for example, a small discrepancy between a numerical value printed in a table and what is supposed to be the same value printed in the text. Some errors, however, are so glaring as to offend common sense.

We present below a detailed analysis of four sample research papers from which the reader may judge the types of errors we counted. Two papers[87,88] from Emory and two from Harvard[92,96] are discussed below as examples of our procedure. Three of these[87,92,96] were eventually retracted wholly[127,A#79,A#170] or in part (A#171), but one[88] as not been

retracted and is still cited as valid (A#131, 132).

The most striking example, in ref. 87, is that in which the authors reported an arresting pedigree depicting a family with a high incidence of an unusual form of heart disease. Inspection of the pedigree, reproduced in our Fig. 1, reveals that a 17-year-old male had 4 children, ages 8, 7, 5, and 4. (It is unlikely that the father's age is a misprint: it appears in the figure and in two places in the text.) Thus the father was 8 or 9 (probably 8) when he impregnated the mother of his first child and 9 or 10 when he impregnated the mother of his second. This bizarre feature of the pedigree, which perhaps could have raised questions about the validity of the entire paper, was apparently not noticed by coauthors or referees, nor were the following unlikely groupings of ages: his sister, brother, and first cousin had their first children at ages 16, 15, and 15, and two women in the preceding generation had their last children at ages 45 and 52.

The summary of the paper gives the urinary taurine levels of the family members with congestive cardiomyopathy as ranging from 411 to 536 mg taurine/g creatinine, but the text and Table 1 of the same paper give for the same measurement 426±45 mg taurine/g creatinine (mean ± standard deviation). Simple inspection suggests that these two sets of numbers cannot simultaneously be valid, and calculation confirms this. Likewise, the summary of the paper states that the urinary taurine levels of the family members with mitral valve prolapse

ranged from 215 to 265 mg taurine/g creatinine, whereas the text and Table 1 give 220±31 mg taurine/g creatinine. These values are central to the thesis of the whole paper. The paper has three numerical discrepancies with ref. 40, and five with refs. 19 and 28, abstracts which appeared before the paper and which deal with the same family. All three abstracts of this work indicated that at the beginning of the study all 46 members of the kindred were alive, which contradicts Figure 1 of ref. 87, in which 3 of the 5 oldest are shown as having died before the beginning of the study.

A more subtle example of a major error is illustrated in Fig. 2, taken from a paper[88] from Emory coauthored by six members of our sample. The figure depicts changes in haemodynamic values in five patients at various time intervals after a surgical procedure. Almost all the data for this paper[88] are contained in three figures which cannot be reconciled with each other. The mean and standard deviation for haemodynamic values in the five patients are given in our Fig. 2 (Fig. 1 in the original paper). It is easy to see that many error bars are the same size, and also that mean SVR (systemic vascular resistance) is almost constant at -160 units (the bottom of the graph) in four serial measurements. These results are especially implausible in the light of data given elsewhere in the paper, where it is shown that SVR and other haemodynamic values for two individual patients vary markedly in the course of time. Moreover, many standard deviations in the

figure we reproduce appear on inspection too small to be reconciled with the widely scattered values shown elsewhere in the paper; this is true of SVR at 0.5 hr, PVR at 1 hr, SWI and SVR at 8 hr, and CO at 12 hr. (Abbreviations are explained in the legend to Table 2.) Indeed simple calculations show that it is virtually impossible to reconcile most of the data in the figure we reproduce with the data in the other two figures of the paper.

The Introduction to this paper[88] indicates that the study of five patients was carried out "in four uncomplicated cases of peritoneovenous shunt insertion, and in one subject who developed signs of cardiac failure and pulmonary edema." But the figure legends indicate that two of the five patients developed pulmonary oedema. Specifically, one of the four patients said to be without complications "developed clinical and radiographic signs of pulmonary edema 45 minutes after the shunt was opened" (Figure 2 legend). The other patient, whose pulmonary oedema was more severe, is described variously as developing this complication "one hour after the shunt was opened" (Figure 3 legend) and "within a few hours" after the shunt was opened (Results section). Furthermore, the statements in the Results section that the changes of the mean shown in Figure 1 after the first half-hour were "very small" and that there was "very little change" appear to be contradicted by the actual maximum changes of the mean -- a factor of two or more in the change from the value at t=0 (corresponding to roughly a 25-60% change in the mean itself) -- for three of the five

parameters in the figure.

There are obvious inconsistencies between this paper[88] and a preceding abstract[30] describing the same five patients. Eight of ten values for mean or standard deviation given in the two publications are inconsistent (Table 2). This includes most of the numerical data in the abstract. Another comparison (Table 2) shows discrepancies in two out of five standard deviations between one abstract[20] and the other[30], and in three of five standard deviations between the abstract[20] and the paper[88]. Thus the two abstracts are inconsistent with each other and with the full-length paper.

In both research papers described above, the major errors are fundamental and call into question the validity of the paper's main conclusions. More typically, however, the errors simply reflect on the care with which the paper had been prepared. This is the case for most of the errors considered below.

Two papers from Harvard are discussed next. The first, reference 92, contains readily detectable errors or discrepancies in Tables 1 and 2; in Figures 3, 6, 9, 10, and 11; and through much of the results section. Some of these are apparent simply on inspection. In Figure 3, for example, most of the error bars are implausibly similar; the same is true of Figure 9. Figure 11 compares two techniques for measuring the same parameter; in Figure 11A, most of the control values from

TTC (triphenyltetrazolium chloride) staining are higher than those derived from histological examination, whereas in Figure 11B, which is based on the same data and should yield the same results, the relationship is reversed. The results with the treated dogs show a similar obvious contradiction of Figure 11A by Figure 11B. Figure 10 (lower right) diagrammatically represents necrotic tissue that is depicted as partially surrounding and touching an ultrasonic crystal transducer; the text, referring to Figure 10, states that the necrotic tissue was "always at least 3 mm" from the crystal.

What are supposed to be exact copies of the same data in tables, figures, and text are actually different in many cases, and it is not possible to determine which are correct. For example, the text gives the preocclusion value for dP/dt in mm Hg/s as 1906±165; Table 1 gives the same value as 1915±192. Similarly, the text gives the percent midmyocardial segment shortening of control dogs after 1 hr occlusion followed by 5 hr reperfusion as -16.7±3.5%, whereas Figure 3 shows the same values as -11.9±2.2%; choosing one of these values over the other markedly affects a test of significance which is described in the text for the first value only (from $p=0.7$ to $p=0.1$). The standard error of the mean of the percent subepicardial segment shortening of the control dogs after 72 hr reperfusion is given in the text as 1.1%; Figure 9 shows the same value as 3.0%.

Among the many errors in ref. 92 are some that bear on

the data which provide support for the main conclusion of the paper -- that treatment with flurbiprofen had beneficial effects on heart function during recovery from experimental coronary artery occlusion in dogs. The text states: "In flurbiprofen-treated dogs a small amount of systolic shortening occurred within 15 min of the infusion of flurbiprofen and persisted for the 3-h period of coronary occlusion". But Figure 9 shows marked passive lengthening at 1.5 hr of occlusion. Also, the legend to Figure 9 labels as significant ($p<.01$) the difference between control and treated dogs (in percent subepicardial segment shortening at 72 hr of reperfusion.) The values actually shown in the figure do not differ significantly ($p>.05$). The text states: "In control dogs there was a small but significant amount of improvement in regional function in the subepicardial segments by the end of 72 h of reperfusion (Fig. 9) . . .". But Figure 9 shows that recovery was about 50% of baseline value, which appears to be a large improvement. Reference 92 contains many errors in addition to those just described (Table 1).

The final example was contributed to a major journal by one of the coauthors about two months after Darsee had admitted to him and another coauthor that he had forged data in another study. At the time the paper[96] was contributed, two senior scientists who were coauthors were carrying out, privately, a "detailed review and analysis" (A#130) of the extent of Darsee's forgery in their laboratory. Thus these two scientists were simultaneously publishing a paper

containing the results of their collaboration with Darsee while serving in private as investigators into the extent of his forgery of data. One author stated (A#159) that he had "total confidence" in this particular paper, which he also said (A#156) "passed special scrutiny." The Ross Committee, asked to review Darsee's work at Harvard and prepare a statement of record, strongly affirmed the coauthors' view that this publication was accurate.

The paper[96] describes the recovery of heart muscle injured reversibly by a brief experimental occlusion of a coronary artery. One measure of heart muscle function (systolic shortening) is said in the legend to Figure 3 not to recover fully "until after 7 days of reperfusion." This description is at least unclear, and we regard it as erroneous, since the figure itself and the data in Table 1 show full recovery at or before 7 days. We present this as an example of an error of very minor significance, but one which we think suggests that the coauthors were not sufficiently careful in examining their own paper before publication. There are small discrepancies between the summary (p. 7152) and the text (p. 7153) with regard to ATP concentration at both 90 min and 72 hr of reperfusion. There is a discrepancy between the error bar for ischemic subendocardium at 15 min of occlusion in Figure 1A and the corresponding number in the summary and text.

The paper also contains discrepancies with another paper[95]

and with an abstract[105], both published previously, and both describing similar or identical studies on similar groups of dogs. The values of maximum negative dp/dt (all between about -4000 and -4300 mm Hg/sec) for one dog in Figure 3 of ref. 96 cannot be reconciled with the values (means from -2248 to -2562; standard deviations from 200 to 316) presumably calculated from this dog plus 6 others and listed in Table 2 of ref. 95. Also, the size of the experimental groups for measurement of ATP and creatinine phosphate at 15 min occlusion and 90 min reperfusion is different in the three publications: 14 or 21 dogs (the text is ambiguous) in ref. 95, 34 or 41 dogs in ref. 96, and 36 dogs in ref. 105. Despite these differences in the numbers of dogs in the experimental groups, the means and standard errors, wherever given, are the same for all three publications. For example, in refs. 95 and 96 (Figures 1 and 2, 15 min occlusion and 90 min reperfusion), there are 16 values and 16 error bars; these are identical in the two papers. Also, in the cases where numerical data are given in the text of the three publications, the numbers after rounding are the same. Such a coincidence of means and standard errors from experimental groups of different sizes is extremely implausible. The paper[96] contains additional errors and discrepancies (Table 1).

Among the 18 research papers, there were as many as 39 errors and discrepancies in a single research paper, with an average of about 12 per paper (Table 1). Of the 22 who were coauthors of a research paper, 19 were coauthors of at least one research paper containing 10 or more errors or

discrepancies.

The analysis of errors and discrepancies given above was based entirely on the coauthors' publications. In contrast, some of the analyses that follow depend on information given in the committee reports.

Retention of data

The reports of the investigating committees showed that in almost all cases none of the coauthors had retained the experimental or clinical data on which the publications were based. This was true of 7 of the 8 Harvard papers, which involved experiments on dogs, and of 9 of the 10 papers from Emory, 7 of which involved human subjects. In the latter case, there were 6 papers for which none of the coauthors had retained either names or other identifying information for the human subjects. Of fourteen in our sample who were coauthors of papers involving human subjects, at least nine failed to retain the list of subjects (Tables 3 and 4).

One paper[25] is particularly interesting: a report on the incidence of mitral valve prolapse in presumably healthy young men. The paper is considered valid (A#132) and is heavily cited (A#181). The coauthors "maintain that the data in the paper are valid although none of them has a copy of the research results. They state that they did the study and oversaw the writing of the paper" (A#132). Although Darsee

was first author of the paper, the Moran report provides no information on his role in the project, and in particular on his opportunities for altering or forging the data; possibly this information is contained in the confidential appendices to the report. Thus the only available support for the validity of this paper is based on a statement attributed by the Moran Committee to the coauthors, given without supporting data or details.

Direct involvement in research

Some coauthors were not directly involved in the published research that bore their own names. "Direct involvement" was defined to mean actual participation in the acquisition and interpretation of data, or the making of an intellectual contribution which was essential to the research. We confined our attention to the 18 research papers. We disregarded a person's rank in the laboratory hierarchy in judging whether an author was directly involved.

Our evidence for the extent of an author's involvement came from the reports of the investigating committees. For example, according to a committee report, one coauthor's role was to encourage Darsee and provide some of the grant support for the research. In another committee report a senior scientist is described as having other duties not allowing him sufficient time for direct participation in the research laboratory and also as being seldom in the laboratory and

having no direct supervisory role. In both these cases we considered that the coauthors had not met our criteria for direct involvement. The available evidence is often limited, and the criteria require a subjective evaluation. Thus it is not clear that another observer would always arrive at the same conclusion.

We consider those not directly involved in the research to be "honorary" coauthors. Honorary authorship was a common practice in our sample, according to our analysis. Of 18 papers, 13 had at least one author judged by us to be honorary. There was a total of 49 authorships (i.e., occurrences of an author's name on one of the 18 research papers); in 16 authorships (33% of the total), involving 6 of the 47 coauthors, the author was judged to have little or no direct involvement (Tables 3 and 4). Those who were honorary authors at least one time published significantly more papers than those who were never honorary authors (3.33±1.63 versus 0.71± 1.42, mean ± standard deviation; $p<.01$ by Student's two-tail t-test). Their higher rate of publication compared with coauthors who were not honorary appeared to depend entirely on their honorary papers: the difference disappeared when an author's honorary papers were not counted (0.67±0.82 versus 0.71±1.42).

The 13 papers with at least one honorary author had more errors and discrepancies than the 5 without an honorary author (15.1±10.2 versus 5.0±6.6; $p=.067$).

Misleading statements

In many ways the most striking of our findings is the discovery that certain research papers (5 of 18) embody statements and data in a manner that misleads the reader about the way experiments were actually carried out. The circumstances were these.

The five papers[53,90-92,94] reported studies done with dogs on the effects of experimental coronary artery occlusion on the heart and on recovery from the effects of occlusion with or without drugs. The papers, all from Harvard, contained statements which we believe the coauthors, at the time of publication, knew or should have known were inaccurate. (When we say that a person "should have known" about an event, we mean that the circumstances were such that a reasonable person using ordinary care would have known about the event.) The published statements that were inaccurate have the effect of making the data and the conclusions look better to the reader than they actually were.

In these papers the assignment of dogs to experimental and control groups is described as formally randomized, as with the statement[90], "Twenty-five minutes after occlusion of the LAD, the remaining dogs were randomized to treated or control groups on the basis of a random odd or even number generated by a computer program." During the subsequent investigation by the

NHLBI panel, however, the coauthors acknowledged that some control data had been reused (A#133) -- in other words, that the control dogs actually included some historical controls, a use which was not mentioned in the published papers. (Historical controls are animals that have been used as controls in a previous study (A#227). The same data recorded in the earlier study are then used again as control data in the later study. The number of dogs for which data were reused is shown in Table 5.)

Clearly the use of historical controls -- even if their use had the beneficial effect of economizing on animal usage and on the investment of time by the experimenter -- produces a less reliable standard of comparison than controls chosen by a randomization procedure at the time of the experiment. For example, failure to randomize as described meant that the surgeon's state of knowledge could systematically affect the outcome of the experiment. Failure to describe accurately the conditions that were actually employed thus concealed from the reader this weakness in the actual experimental design.

The senior scientist, who was a coauthor of three of the five papers, subsequently stated (A#134): "In order to economize animal usage, as well as professional and technical time, control animals which had previously been randomized were sometimes used to provide data for more than a single experiment . . . . we saw no reason not to allow [Darsee] to use his randomized controls in more than a single

experiment." The other coauthor of all five papers acknowledged the reuse of data, but stated (A#135) that Darsee had misled him and the other coauthor about the number of dogs for which data were reused and about the number of experiments in which the data for any one dog were reused. As far as we know, neither author has ever publicly claimed that at the time of publication he was unaware of the inaccuracy of the published statements about experimental conditions or that he did not sanction the use of historical controls.

Reuse of data was approved by one of these coauthors in at least one paper not involving Darsee (mentioned in ref. A#133), and both coauthors have subsequently defended these practices. One coauthor, commenting on a study not involving Darsee that was carried out in his laboratory, acknowledged (A#135) that control data were reused and that the reuse was not described on publication: "Although the use [of some control data from another study] was not specified in the text [of a study not involving Darsee], the use did not bias the results of the study in any way." The other coauthor defended the practice of sharing control results said to be randomized by stating (A#134) that "the practice of not specifying in the Methods section that a control dog had been used in more than one study, while certainly not ideal, is common," and he also stated (A#134) that "the practice of utilizing controls in more than a single experiment is not uncommon and, <u>although it should be</u>, it is not always spelled out in published papers" (emphasis added). To the extent that a control group used in a

physiological experiment provides merely a yardstick by which the response of an experimental group may be judged, the use of historical controls is not strictly impermissable. We note, however, that such a procedure is inconsistent with the statement in the paper quoted above that "after occlusion of the LAD, the remaining dogs [i.e., those surviving the occlusion] were randomized to treated or control groups on the basis of a random odd or even number generated by a computer program"; moreover, the statement that was published depicts an experimental procedure superior to that which was actually followed.

As pointed out in the NHLBI panel report, the use of historical controls was associated in one instance with a statistically significant bias ($p<0.03$) that had the effect of exaggerating the efficacy of the drug under study. The panel considered it likely that Darsee alone was responsible for the bias in selection of reused data, and we concur; accordingly we have not attributed this bias to the coauthors or included it in Tables 3 or 6. Nevertheless, the bias in selection of data was made possible in part by the practice, which was sanctioned by the two coauthors, of using historical controls.

It would appear from the subsequent statements of the two coauthors that they knew of and authorized the use of historical controls in the laboratory, as well as the failure to describe that use on publication. Since the described randomization procedure in any study was supposed to apply to

all dogs in that study, but in fact did not apply to the historical controls when their data were reused, we conclude that the coauthors knew or should have known at the time of publication that the papers contained an inaccurate description of experimental procedures.

The misleading description went beyond the supposedly random method of assigning treated and control dogs. Specifically, the experiments reported in the five papers included procedures, such as arterial catheterization and administration of certain drugs or dyes, that differ from one study to another (Table 7). As a result, most of the historical controls had been subjected to procedures in an earlier study that were different from the procedures described for the later studies. Thus the published descriptions of these later procedures, which are described explicitly and in detail as applying to all dogs, do not apply in some respects to the historical controls.

One author, writing to the NIH about the sharing of results, stated (A#134) that "the experimental conditions were similar" for the studies in question. The other referred (A#135) to studies (not involving Darsee) in which "experimental conditions are the same"; he has also indicated in correspondence with us that the procedures described in the five papers were similar enough in some cases so that the use of historical controls was justified. Differences in experimental conditions are summarized in Table 7. Some might consider that in most cases shown in Table 7, the data from

dogs in one study would be unsuitable as reused data for another study, but the facts enabling readers to form a judgment on this point were not given in any of the five papers. We believe that the coauthors knew or should have known at the time of publication that the descriptions of experimental procedures did not apply to the reused data obtained in previous studies and that these descriptions were therefore inaccurate.

A final point may be related to reuse of data: one of the five papers states[91] that "all dogs [were subjected to an experimental procedure, which is described]. Ventricular fibrillation or standstill was then induced by a barbiturate overdose, and the hearts were excised rapidly and placed on ice." Yet the coauthors have said (cited in ref. A#133) that some of the same dogs were subsequently used for the 4-hour and 6-hour occlusions in other experiments. As the NHLBI panel report politely observes (A#133), "This would appear to be impossible." We have assumed that the two coauthors in our sample were unaware of this anomaly at the time of publication.

The inaccurate statements noted above appeared in five papers coauthored by two scientists in our sample (Table 3). In addition to these five papers, there were two others (refs. 95 and 96; see next section) containing statements that were inaccurate or misleading. In total there were 18 instances of statements made by 6 coauthors in our sample under conditions where they knew or should have known that the statements were

inaccurate or misleading.

Unacknowledged republication

We found several instances in which the same data were republished without adequate reference to that circumstance. The most striking example was a paper[95] containing data later published almost in their entirety as a major part of a subsequent paper[96] which appeared in a different journal. Although the second paper referred to the first by means of an index number in a list of references, there is no explicit statement in the second paper that the new data are a continuation of an experiment already published.

The subject matter of both papers concerns the recovery of heart muscle injured by temporary surgical occlusion of a coronary artery in dogs, described by the authors as a model with "important clinical implications for patients" with certain kinds of heart disease. The first paper states that "Seven dogs were also subjected to 15 min of LAD occlusion followed by 3 days of reperfusion", which is not false but which does not reveal that the same 7 dogs were actually reperfused for 14 days. The references in the second paper to the first would not suggest to those who were not familiar with the first that the research about to be described was a continuation of an experiment already published and that all the data had been collected (but perhaps not analysed) before

the first had been submitted for publication. In the second paper's opening paragraph, which is reproduced here in Table 8, neither citation of the first paper provides a hint that some of the same data are being republished and that some of the same animals were used.

From the evidence available to us, there is no way of being certain whether this way of dealing with the references to earlier work was intended to conceal the connection between the two papers or whether it was the consequence of the clumsy or careless use of language.

On neither paper was Darsee the first or last author, and it seems that the coauthors should have known of the statements that were misleading or inaccurate. For the journal to which the second paper[96] was submitted (Proceedings of the National Academy of Sciences), a contributing author customarily bypasses the refereeing system, which might have discovered the unacknowledged republication of data; instead the contributing author assumes "responsibility for the propriety and scientific standards" (A#202) of the manuscript he submits.

A common example of republication involves abstracts of research submitted to scientific societies. Of the 88 abstracts, only 47 were published only once; all the rest involved duplicate or triplicate publication. This practice is apparently followed in many fields of biomedical science, where an abstract may be prepared for one meeting of a society and

then submitted, perhaps in an amended form, to a second meeting. We nevertheless suspect that the extent of this practice may not be widely appreciated. For the purposes of the present report we have disregarded this practice.

We note that there were four pairs of Darsee abstracts in which the texts were nearly identical, but whose titles had apparently been changed. For example, an abstract [77] entitled "Can tetrazolium stains reliably identify myocardial infarction prior to definite histologic necrosis?" was also presented [73] under the title "Early pathologic detection of acute myocardial infarction." Sometimes the alteration was made by substitution of synonyms [72,80]: "Persistent myocardial abnormalities following brief periods of temporary coronary occlusion not associated with necrosis" was changed to "Prolonged metabolic, functional, and ultrastructural abnormalities following transient myocardial ischemia without infarction." In each pair the authors are the same, and the text and numerical values are identical through most of the abstract. The publication dates were within about a month of one another, as were the scheduled presentation dates. Two coauthors of some of the abstracts have stated privately that the titles of certain abstracts were changed to make them acceptable for oral presentation before the different societies to which they were submitted. However, we consider that the change in title has the effect of making the pair of abstracts, if listed only by title, appear different to the reader, whereas in fact the contents are almost the same.

There were 5 instances of unacknowledged republication of data in our sample. They appeared in publications, one from Emory and four from Harvard, having a total of 8 coauthors (Tables 3 and 4).

Practices departing from accepted standards

We were surprised to find that there is apparently no generally accepted written code of conduct for scientific research (A#243,245-250,252,253,260,261). We think that there is nonetheless an unwritten code of which scientists are aware and to which they profess adherence. We have based the analysis that follows on standards that we believe are generally accepted -- accepted in the sense that they are generally understood, not necessarily the sense that they are always complied with. Our belief was supported by extensive discussions with many colleagues; among those who read our report, almost no one said he thought the practices described here were acceptable. Both in our description of events, given above, and in the interpretation that follows, we realize that there is necessarily a subjective component, and that readers of our report may differ in their view of what is acceptable and what is not.

In Table 6 we describe ten categories of practices that were chosen to fit the practices we observed in our sample; a different set of practices would probably be more appropriate

for other studies. We found it useful to divide these practices into two groups. Those classified as type A could be explained simply by carelessness or excessive haste, while those of type B appeared to be more serious in some sense. We do not mean to suggest that certain instances of type B practices may not be due solely to carelessness and haste; undoubtedly this is so in some cases. Furthermore, we recognize that the distinction between type A and type B practices is necessarily subjective and that other observers might classify them differently.

We emphasize that the instances listed in Table 3 were based on evidence we considered compelling, but not necessarily conclusive in each case. We did not include instances which we think probably occurred, but for which the evidence was only suggestive. Thus the frequency of type A and B practices reported here is probably an understatement of the actual frequency in our sample. We consider our data adequate to estimate the minimum frequency of practices in our sample; we recommend that our conclusions not be used to hold individuals accountable for particular practices.

Frequency of practices

Of the 47 coauthors in our sample, 31 (66%) had, according to our analysis, engaged in practices of type A -- a result that is the same whether or not honorary coauthorships are included (Table 3). Furthermore, 13 (28%) had engaged in type B practices. Only 12 of the 47 were not found to have

engaged in practices of either type; of these 12, 10 had not authored a research paper.

Some scientists apparently do not regard abstracts as publications, even though they are archived and cited and are sometimes used to establish a claim to priority (A#254). If the 88 abstracts and 3 chapters are excluded from consideration, the frequency of type A and B practices is about the same or higher. Of the 22 scientists in the sample who were coauthors of the 18 research papers, 21 (96%) engaged in practices of type A, and 6 (27%) engaged in practices of type B.

The number of type A or type B practices by an individual appears to increase with the number of his publications (see Table 3), a result that might be anticipated on several grounds. Interestingly, authors with more publications tended to have more practices (of type 3-10) per paper than those with fewer publications ($r=0.51$; $0.01<p<0.05$; data not shown).

Representativeness

The evidence presented here indicates a surprisingly high frequency of behavior that departs from accepted standards in the sample that was examined. One possible explanation is the association of these scientists with Darsee: he may have induced them to behave in ways not typical of them. A second possibility is that the scientists in this sample are not representative of their field. A third possibility is that

practices that depart from accepted standards may be unusually common in the specialty in which these scientists worked, in clinical research generally, or in research carried out at medical schools or by physicians. (Our sample was mainly physicians, and the research was carried out exclusively in medical schools.) A fourth and disturbing possibility is that practices that depart from accepted standards may be more common among biomedical scientists than is currently realized.

Other studies will be needed to determine whether our results are representative of other groups of scientists.

Discussion

The end product of scientific research is information communicated to other scientists or used in some application. If that information is degraded -- whether through carelessness or deceit -- the effect is a waste of resources, which could easily be in the hundreds of millions of dollars annually in the U.S., and a slowing of scientific progress, both in basic and applied research.

When there is professional misconduct by biomedical scientists, this presumably delays the discoveries that lead to the prevention or treatment of disease. Since the amount of the misconduct is unknown, one cannot estimate the size of the resulting delays.

It is obvious that the practices described here have at least the potential for interfering with the accuracy of the scientific literature or for harming the scientific enterprise in some other way. We also think that the instances described above represent departures from the generally accepted standards of scientific research. We recognize that the decision as to what constitutes a departure from generally accepted standards is necessarily subjective. Presumably some scientists have views different from those expressed here. We regard it as a matter of the greatest importance that these differences be discussed publicly.

We summarize briefly some effects of each of the practices we have described.

The abundance of errors that we observed in our sample obviously interferes with the accuracy of the scientific literature. Every scientist has, at a minimum, an obligation to ensure that what is published under his name is accurate.

There is a notable absence of explicit, published standards for the retention of data. We believe that it would be generally accepted that data should remain accessible for as long as questions and criticisms are likely to be raised about what has been published. For most of the papers in our sample, data were not available. We used a conservative criterion: we disregarded all cases not involving human subjects. We counted as departures from generally accepted standards only those

cases where data involving human subjects had not been retained. We suggest that a reasonable minimum requirement is that each author of a paper involving human subjects should retain at least a list of the subjects' names or the hospital's identification numbers of patients. This information would permit access to hospital records and thus in many cases allow reconstruction of data in the paper.

The reader may ask: What harm is done by honorary authorship? Indeed some of our colleagues have argued that the custom of routinely placing the name of a senior scientist, usually the head of the laboratory, on a paper -- regardless of his contribution -- is widely followed and does no harm. We disagree, as have others (A#151,172,239,241,251,257,261): honorary authorships falsify the assignment of responsibility for published research and increase the likelihood that inaccurate data will be published. The honorary author is in a poor position to judge the validity of the work, yet he often lends a prestige that may lull other coauthors, the reviewers, or the readers into uncritical and inappropriate acceptance.

Similarly, the harm done by unacknowledged republication of data may not be immediately obvious. But the practice disguises the true relationship between published papers and yields no benefits, except possibly a personal benefit to the authors involved.

In some ways the most disturbing of our findings is the

discovery that certain of the publications (five out of 18) embody statements and data in a manner that misleads the reader about the way experiments were actually carried out.

The ten practices discussed in this report have in common the following insidious feature: the published research may be defective, but it appears sound.

Causes of misconduct

Whether or not our data are in fact representative, it is obviously important to ask what causes departures from accepted standards of research.

Questionable practices among scientists may perhaps be related in some way to their previous behavior in school. In most studies roughly 30-60% of students are found to cheat in high school (A#175,176), college (A#147,175,177,178,200), graduate school (A#148,200), and medical school (A#147). It seems unlikely that those who successfully cheat as students are somehow excluded from a career in research, or that they enter research but turn over a new leaf.

Table 3 lists by author the 219 instances, 180 of type A and 39 of type B, in which there were departures from accepted standards of research. One motive could account for most of these -- the desire to maximize the number of papers on which one is an author.

Extraordinary emphasis is commonly placed on sheer quantity of publications (A#149,193-195,198,201,235,238,242, 251,255-260). The director of one of the world's leading research institutions sent a signed, official memorandum to junior colleagues proposing incentives for publication: "There is no demand that these be literary masterpieces in first line journals; journeyman works for publication in second, third or fourth line archival publications will be quite satisfactory. Upon proper completion and submission of [two] manuscripts, [a technician]'s appointment will be extended to April 1, 1984. During that time it is expected that an additional manuscript on [subject] will be completed and submitted. If so, the period of employment will be extended an additional three months and again an additional manuscript on [subject] is an anticipated result of the extended employment." On another occasion a director of research said in all seriousness, "No research that results in a single paper is worthwhile."

There are several indications in our data that this sort of pressure may have led to many of the departures from accepted standards in Table 3. Coauthors who published more papers tended to have more departures (of type 3-10) per paper than coauthors who published less. Coauthors who were honorary one or more times published more papers than those who were never honorary -- but only if their honorary papers were counted. Papers with honorary coauthors had more errors than those with none.

Those promoting Darsee or recommending his promotions were well aware that his publications were unusually numerous (A#134). But apparently none of these supporters, who were also his coauthors, made the effort that would have shown that many of the papers had serious flaws. How often does this occur?

Despite the lack of conclusive evidence, it seems likely that the extreme emphasis commonly placed on producing large numbers of publications, regardless of accuracy, quality, or probable usefulness, was a major factor in the origin of the practices described here. Also, the extreme competition for research grants and the phenomenon of team science may also contribute to the genesis of practices that depart from accepted standards of research (A#234,235,237).

The absence of a written code of conduct for scientific research may facilitate practices that depart from accepted standards. The editor of a leading biomedical journal told us after reading a draft of our report that we failed to make allowance for the "subtleties and ambiguities of real life"; no one, he said, lives by what we call the accepted standards of research. Is this true?

The obstacles we encountered -- not described here -- in completing and publishing this report may account for the virtual absence in the published literature of empirical studies

on the frequency of misconduct among scientists.

Some scientists are seemingly unwilling even to recognize the existence of misconduct. Following a recent episode of widespread cheating among several thousand new medical school graduates (A#180), the American Medical Association's vice president for medical, educational, and scientific affairs described (A#180) honor and non-cheating as "important acts of faith." On another occasion a distinguished scientist, chosen as an auditor in a case of alleged fraud, deferred his investigation for almost six months and then withdrew. When the charges of fraud were later shown to be valid, he said (A#153), "I just found it hard to believe that [the head of the laboratory] had engaged in any hanky panky. I saw the validity of the complaints in the [accusing] letter, but my feeling by and large was that people don't falsify data and don't plagiarize things."

It appears that most scientists have not given much thought to the practical or ethical implications of the practices described here; one scientist referred to these practices as "little scientific white lies." Indeed the professional success, if not survival, of many scientists may depend on their continuing to engage in these practices, which often earn them substantial rewards. Many see these practices as privately sanctioned by those in authority.

There is currently almost no institutional support for

studies on the frequency of scientific misconduct. Based on the comments of other scientists, as well as our own personal experience, we think that if such studies are begun, attempts may made to interrupt or censor them; if completed, they may be unacceptable to the editors or publishers of scientific journals. As a consequence, even those scientists who regard the subject as interesting and important may avoid working on it.

Honorary authorship and the research czar

The essence of honorary authorship, as we have defined it previously, is the claiming of scientific credit where virtually no scientific contribution has been made. Often this occurs when a senior scientist, who controls the funds for a research group that he directs, obtains credit by authorship for scientific contributions that are in fact minimal. The practice appears to be widely tolerated by scientists. This tolerance has thus created a new class of scientific bureaucrat -- the research czar -- who is a prominent and influential (although often feared) member of the scientific community. In such cases, the chief of the laboratory profits from the convention that his name appears on publications from his laboratory, even though his scientific contributions may have been minimal, and his real activities were the acquisition of funds and certain administrative functions.

This misattribution of scientific credit and the widespread

tolerance of the practice have a number of harmful effects that go beyond the unfairness of taking scientific credit where it is not due. First, the combination of power and remoteness in the chief of the laboratory may lead him to make a poor choice of research projects, to insist that they be carried to completion, perhaps with inappropriate methods, and finally to make an incorrect judgment of the accuracy of the research results. Nevertheless, as we have mentioned, his high position lends a prestige that may lull or intimidate other coauthors, the reviewers, or the readers into uncritical and inappropriate acceptance of the work as valid. Second, sheer numbers of publications accumulated by a laboratory chief become a measure of his success as a scientist, and he is likely to apply the same criterion in judging other scientists in his own laboratory and elsewhere, with the harmful consequences we have discussed previously. Third, the illusion is created that the chief is doing research, when in reality he is not. The so-called Principal Investigator (as the holder of research grants is formally designated) may in these cases not really be an investigator at all. We emphasize that the above comments apply solely to research that is done by a team of scientists led by a scientist-administrator not actually involved in the research.

Those chiefs who are skillful at acquiring authority may value and reward conformity over adherence to sound research practices. One writer has used the term "hired hand research (A#240)". When the chief's expectations differ from the facts, the junior scientist on the team faces a hard choice.

It is probably no accident that many of the discovered instances of fraud have occurred in large research groups. Such groups offer a perfect setting for a forger, as well as an environment that permits, or even encourages, type A and type B practices.

A strongly hierarchical structure may be appropriate for military and business organizations and for certain types of applied research, but it is seems uniquely inappropriate for basic research and for most research in the biomedical sciences. Indeed the rigid hierarchy characteristic of Soviet science has been given as a reason for the fact that their biomedical science lags far behind that of the West (A#263).

A test of the system

How well does the system of scientific self-regulation work? No one knows. In some cases of outright forgery of data, the customary methods of checking published research have worked well, and the problem was recognized and corrected quickly (A#229-232). It is especially likely that forgeries of important results will rapidly be detected. But there is no reason to believe that scientific self-regulation consistently detects or discourages the less flagrant and presumably more frequent departures from accepted standards of research like those discussed in this report. Indeed our observations suggest that the system of self-regulation may be inadequate.

In the group of papers analyzed in this report, it is striking that the errors were numerous, that some call into question the validity of the work, and that the errors appear to have escaped the notice of coauthors and referees. Initially we chose our sample because of the large amount of information in the reports of the investigating committees. In the event, however, it was remarkable how much information was obtained simply by a careful reading of the coauthors' publications. Indeed this forms the basis for our proposal, described below, to audit a random sample of published papers.

Although the coauthors in the present study may not be statistically representative of biomedical scientists in general, this is less likely for the referees and others who were responsible for checking the work. That these papers were published in leading biomedical journals despite their overt and covert defects suggests that the usual mechanisms of peer review may be less effective than is commonly believed in protecting the accuracy of the scientific literature. The function of the refereeing system as at present practiced, is not merely to authenticate published material but also to endow the authors of what is published with the implicit approval of their peers. The deficiencies of the system have been remarked upon on several occasions (A#209-211,244). Some of our findings -- error bars of uniform height, for example, or genealogies that make no sense -- seem to us to reinforce these doubts.

In the episode of forgery that initiated this report, the response of the scientific community proved inadequate even though there was clear evidence of serious professional misconduct (A#155,160,161). Investigating committees were appointed at both medical schools, but the disappointing performance of these committees suggests that this mechanism for studying alleged misconduct -- a committee chosen by the dean largely from within his own institution -- may not be effective.

After the confession to forgery of data on May 22, 1981, Harvard Medical School immediately and privately withdrew its offer of an assistant professorship to Darsee. Harvard Medical School acted rapidly to protect itself, but over the next five months it took little action to protect public funds, other institutions, or other collaborators. Darsee was allowed to continue on the "Models" study, a costly project involving several other institutions and many scientists. The NIH, which was supporting research projects on which Darsee was working and had paid for his postdoctoral fellowship, was first informed (A#133) of Darsee's forgery on October 29. A senior scientist collaborating with Darsee in another research group at Harvard was not notified until November (A#130,133).

The dean of Harvard Medical School, centrally involved in the initial decision not to disclose the problem publicly, subsequently appointed a committee to investigate both the incident itself and the institutional response within Harvard. This appears to have created a conflict of interest. The Ross

Committee, appointed by the dean, examined Darsee's work at Harvard. Five of the eight Ross Committee members were on the faculty at Harvard Medical School or other Harvard schools, and (though not stated in the report of the committee) the chairman of the committee was a member of the Harvard Board of Overseers.

These defects, however, were soon overshadowed by the magnitude of the Ross committee's erroneous conclusions. The committee stated, among other things, that a "systematic review of all work involving Dr. Darsee has been conducted by [the two senior scientists in charge of the laboratory]. Primary data have been reviewed and Dr. Darsee's co-workers have been questioned" (A#130). But the committee failed to interview the junior coworkers (A#156,160). This omission proved unfortunate, since several were highly suspicious that Darsee was systematically faking a major part of his research (A#156,160), a view they had previously communicated to both of the senior scientists in charge of the laboratory (A#156,160). The Ross Committee report continues: "The committee is convinced that [the senior scientists] have documented the extent of the irregularities and properly concluded that the previously published work from the laboratory in which Dr. Darsee was a member of the group is accurate. None of the work under suspicion has been published. The papers and abstracts containing fabricated data have been withdrawn. This incident has resulted in a strong reaffirmation of the laboratory rules concerning data collection, review and retention" (A#130).

This last statement in retrospect seems wildly optimistic, even if all the data were available. But a few months later the NHLBI panel (A#133) found insufficient primary data to verify the work for seven of the eight Harvard papers containing original experimental data. Either the primary data had not been reviewed by the senior scientists as stated, or they were reviewed and discarded. Either possibility seems inconsistent with the Ross Committee's finding (A#130) of a "strong reaffirmation of the laboratory rules concerning data. . retention." Eventually all but one[93] of the papers coauthored by Darsee and deemed "accurate" by the Ross Committee were retracted wholly[53,90-92,94] or in part[95,96,124,125].

The Ross Committee report contains many other obvious and remarkable deficiencies, but it was endorsed over to the public by the dean with great thanks to the committee for the "extraordinary service that they have rendered to the Harvard Medical School, to academic medicine and to society at large" (A#129). An editorial in the New York Times called the Harvard response "a model of obfuscation" (A#154).

The Moran Committee, appointed by the dean and composed exclusively of faculty members at Emory University School of Medicine, failed to obtain relevant data, used insufficient care in evaluating data and withheld data that might have enabled a reader to judge the validity of the committee's conclusions. The committee failed during its year of investigation to

discover the existence of 18 Emory abstracts (A#131); even after the release of an addendum (A#132), six abstracts[8,21,34-36,39] -- all listed in the ISI Source Index under Darsee's name -- were still missing from the committee report. Coauthors frequently gave the committee an evaluation of their own data instead of providing the actual data (A#131). This proved unsound: one paper[24] declared valid in the committee's report was later judged invalid in an addendum. Also, the committee failed to interview some coauthors of the papers in question (A#131). It even failed to discover the existence of one coauthor who worked at Emory.

Because the Moran committee members used insufficient care and relied on evaluations by the coauthors, they mistakenly accepted the validity of ref. 88, despite its gross defects, which we have discussed here. An external university committee afterwards praised the Moran committee for the thoroughness of its work (A#166), and the then dean of the Emory medical school has more recently published a note (A#183) affirming the validity of the two abstracts[20,30] associated with the paper[88]. Although the paper contains recommendations concerning the drug treatment of patients, it has not been retracted and indeed was cited on three occasions (A#167-169) in 1982.

Finally, much of the data of the Moran report is contained in some 69 appendices of an unusual nature. These appendices, cited frequently through the report for substantiation, are all secret! The reader is unable to judge whether the committee's

conclusions are valid. Indeed, there is an element of secrecy in the whole report (even though it was given to certain members of the press): it was unpublished for three years, and despite vigorous solicitation, we were unable to obtain a copy from the office of the dean at Emory. (We finally obtained one, without appendices, from a friendly science journalist.)

We suggest that the standard procedure for dealing with suspected fraud should be the appointment of a committee most members of which are unaffiliated with the institution in which the investigation is carried out, and with its members preferably not chosen by the dean. The complete findings of the committee should be made public. The committee should start with the premise that, after a single forgery has been established, all other work of that scientist should be considered invalid until proved otherwise.

Journal policies do not always encourage correction of error. The editor of a major biomedical journal explained (A#206) to some of the coauthors of four questionable papers published in his journal why the retractions they wished him to print should be published elsewhere: "I received your letter . . . concerning the fact that [certain data in the four papers are] potentially fraudulent. Upon reflection, I do not believe THE JOURNAL should publish any note of this, since it hardly seems a meritorious issue that justifies breaking our tradition of publishing only original scientific work." This journal is not alone in its reluctance to publish retractions (A#208,234).

In the course of preparing this report, we had many private discussions with senior scientists. A large fraction of them have encountered instances of serious professional misconduct in their own laboratories, but have taken little or no action for fear of the legal or social consequences. This suggests to us that serious professional misconduct may be more common than is currently realized, and further suggests that the institutional mechanisms for dealing with this problem are inadequate.

The same pressures that lead susceptible scientists to engage in misconduct may have pervasive, harmful effects on the way that almost all scientists plan and carry out their work. In a conversation with one of us, a scientist at one of the nation's leading medical schools sheepishly referred to much of his own work as "semi-meaningful research," which he explained as research largely lacking in importance, but chosen because it could be counted on to produce publishable results. It is our impression that this practice is common and that it often has institutional encouragement and reward.

Effects of misconduct

We have found it convenient to divide the practices discussed here tentatively into two groups, type A and type B. Our division is shown in Table 6, and the number of instances is shown in Table 4. At this point we should emphasize that this

distinction is not unambiguous and that we do not intend to signify that any particular instance of type B may not be capable of explanation by carelessness or haste alone. Nevertheless, the practices listed as type B appear in general to go beyond carelessness or haste.

We do not discuss at all in this report the practices that we may refer to as type C, namely, those involving wholesale forgery of data, such as that which led to the selection of our sample of scientists. The rare examples of type C practices that have come to light seem to elicit from scientists, journalists and laymen alike a reaction that is strong, critical and emotional. Many writers and some scientists appear to feel that type C practices threaten the foundations of science. We suggest that this may not be true: in many instances of type C practices that have been described, relatively little long-lasting harm appears to have been done (A#149,229-232). Furthermore, because these instances of type C behaviour appear to be rare, it seems reasonable to speculate that in their aggregate effect, type C practices do little harm to science.

Many readers of our report have argued that it is the type B practices that are urgently in need of attention, perhaps even to the exclusion of type A practices at the present time, because the type B practices are in some sense "worse" than the type A practices. This point, however, strikes us as weak. It is simply not known whether individual instances of type A practices do more or less harm to science than individual

instances of type B practices. Moreover, it seems reasonably likely that practices of type A are more common than those of type B. This suggests a surprising conclusion: in their aggregate effect, type A practices may be much more harmful. We therefore recommend strongly that the frequency and effects of type A practices receive more attention.

The most important questions pertain not to type C practices, in our opinion, but to type A and type B: how common are they, and how much harm do they do? We suggest that the scientific community can no longer afford to ignore these questions.

How many publications are not supported by data?

The results of our study naturally raise the question: what fraction of papers in the biomedical sciences are not supported by primary data at the time of publication? The question cannot be answered with the papers in our sample, since most are known to be partial or complete forgeries and thus are atypical.

We suggest a study to answer the question: close examination of a random sample of published papers. Probability of selection could be uniform or could be weighted by a variable such as number of authors or cost of research. Papers would be examined by either of two methods: careful examination of the

paper itself for errors and discrepancies by appropriate experts following an established procedure ("external audit"), or examination of the primary data on which the paper is based, which would of course require access to laboratory records ("internal audit"). The advantages of the external audit are its noninvasive character and presumably lower cost; the advantage of the internal audit is its completeness.

What would this cost? This depends on the precision required and on the actual fraction of defective papers. We assume for the sake of illustration that an external audit could be done for about 3% of the cost of the research on which a paper is based, and an internal audit for about 10%. If the actual fraction of papers that could not be substantiated were as high as 15% and if the answer were required with a standard error of 7%, about 25 papers would have to be examined, which in the case of internal audits would altogether cost only about 2.5 times as much as one average paper. If, on the other hand, auditing 100 papers failed to disclose a single paper that could not be substantiated, this would show with $p=0.05$ that the actual fraction of papers unsupported by primary data at the time of publication is less than 3%.

In addition to the financial costs of examining the practices of scientists, there are other costs which may be more serious in the long run. Examination of scientific practices could cause unwarranted harm to individual scientists. Systematic examination of scientific practices might even weaken

the fabric of trust that is essential to the functioning of science. Of all human endeavors, science is one of the most successful -- prodigious in benefits, relatively low in cost. But science, vulnerable to abuse from within by its practitioners, is perhaps even more vulnerable to harm by regulation, and at some point the cost of further regulation will outweigh the benefits.

Among the scientists who read our report, the clear majority praised it, but some expressed strong reservations. They pointed out that our sample may well be biased, since it was not selected randomly, and indeed the scientists in our sample were chosen through association with a known forger of data. For these reasons the frequency of type A and B practices we found may not be representative of that in biomedical science as a whole. Some pointed out that the criteria we use are subjective and that others might arrive at different conclusions. All agreed that an accurate assessment of the frequency and harm of these practices will require a study that is better than the one presented here. We agree with all these comments, particularly with the need for a better study.

What should be done? A proposal

The frequency of misconduct among biomedical scientists is still an open question. Three science administrators (A#179,180,262) have previously given figures indicating a frequency of detected scientific misconduct of about 0.1%. In

our sample we found a frequency of almost 70% -- seven hundredfold higher than the previous estimate. Even if only the type B practices are considered, the figure is almost 30% -- still several hundredfold higher than the previous estimate. The figures on frequency of misconduct that emerge from our study are clearly not tolerable and, if widespread, pose a profound threat to the integrity of scientific communication.

The data and discussion above point to a problem that is substantial, but whose exact size is unknown. However, we feel it is not too soon for scientists to start considering solutions to the problem, regardless of its exact extent. Our data clearly indicate the need for scientists, especially the leaders of science, to pay more attention to the validity of published research and the accuracy of the scientific literature.

The practice of profiting from research that one has not actually done appears to be centrally involved in the origin of many practices we observed in this study. Accordingly, it seems reasonable to consider ways of reasserting the principle that scientific credit be given only for scientific work actually accomplished. No single approach is likely to solve this complex problem. Some journals are now requiring from authors a statement that defines their participation in submitted papers. We suggest another possible approach: decreased scientific support for large research groups directed by a single chief, and correspondingly increased support for individual scientists. This would not prevent the individual scientists from joining

together in teams wherever that seemed appropriate for them and advantageous for the investigation of particular scientific problems.

Scientists have to an unusual degree been entrusted with the regulation of their own professional activities. Self-regulation is a privilege that must be exercised vigorously and wisely, or it may be lost.

References

1. Notre Dame Sci. Q. 7, no. 2, 14-17 (1969).

2. Notre Dame Sci. Q. 8, no. 1, 9-11 (1969).

3. Circulation 56, III-218 (1977).

4. South. med. J. 70, 1249 (1977).

5. Archs intern. Med. 137, 1639 (1977).

6. Ann. intern. Med. 89, 867-870 (1978).

7. Clin. Res. 26, 3A (1978).

8. Clin. Res. 26, 226A (1978).

9. Clin. Res. 26, 226A (1978).

10. Chest 74, 334-335 (1978).

11. Chest 74, 334 (1978).

12. Circulation 58, II-237 (1978).

13. Circulation 58, II-235 (1978).

14. Circulation 58, II-229 (1978).

15. Circulation 58, II-174 (1978).

16. Circulation 58, II-42 (1978).

17. Circulation 58, II-23 (1978).

18. Clin. Res. 26, 773A (1978).

19. Clin. Res. 26, 762A (1978).

20. Clin. Res. 26, 759A (1978).

21. Clin. Res. 26, 745A (1978).

22. New Engl. J. Med. 298, 221-222 (1978).

23. Am. J. Med. 64, 1089 (1978).

24. Circulation 59, 492-497 (1979).

25. Circulation 59, 619-622 (1979).

26. New Engl. J. Med. 300, 877-882 (1979).

27. Am. J. Cardiol. 43, 367 (1979).

28. Clin. Res. 27, 274A (1979).

29. Clin. Res. 27, 265A (1979).

30. Clin. Res. 27, 264A (1979).

31. Clin Res. 27, 230A (1979).

32. Clin. Res. 27, 161A (1979).

33. Abstract of report presented in annual Lloyd Hyde Medical House Staff and Fellows Research Program, Emory University School of Medicine (1979).

34. Clin. Res. 27, 573A (1979).

35. Clin. Res. 27, 561A (1979).

36. Clin. Res. 27, 561A (1979).

37. Chest 76, 371 (1979).

38. Chest 76, 350 (1979).

39. PACE 2, A43 (1979).

40. Circulation 60, II-157 (1979).

41. Circulation 60, II-149 (1979).

42. Circulation 60, II-124 (1979).

43. Circulation 60, II-90 (1979).

44. Circulation 60, II-41 (1979).

45. Circulation 60, II-38 (1979).

46. Clin. Res. 27, 726A (1979).

47. New Engl. J. Med. 301, 443 (1979).

48. South. med. J. 72, 174-180 (1979).

49. Circulation 60, II-252 (1979).

50. Ann. intern. Med. 92, 735-741 (1980).

51. Am. J. Cardiol. 46, 607-612 (1980).

52. Am. J. Cardiol. 46, 613-618 (1980).

53. Am. J. Cardiol. 46, 800-806 (1980).

54. Am. J. Cardiol. 45, 492 (1980).

55. Am. J. Cardiol. 45, 476 (1980).

56. Am. J. Cardiol. 45, 395 (1980).

57. Clin. Res. 28, 257A (1980).

58. Clin. Res. 28, 163A (1980).

59. Clin. Res. 28, 163A (1980).

60. Clin. Res. 28, 653A (1980).

61. Chest 78, 513 (1980).

62. Chest 78, 513 (1980).

63. Circulation 62, III-318 (1980).

64. Circulation 62, III-317 (1980).

65. Circulation 62, III-177 (1980).

66. Am. J. Cardiol. 45, 394 (1980).

67. Fedn Proc. 39, 1112 (1980).

68. Fedn Proc. 39, 398 (1980).

69. Clin. Res. 28, 538A (1980).

70. Clin. Res. 28, 163A (1980).

71. Clin. Res. 28, 174A (1980).

72. Clin. Res. 28, 615A (1980).

73. Clin. Res. 28, 614A (1980).

74. Clin. Res. 28, 610A (1980).

75. Clin. Res. 28, 609A (1980).

76. Circulation 62, III-314 (1980).

77. Circulation 62, III-246 (1980).

78. Circulation 62, III-239 (1980).

79. Circulation 62, III-196 (1980).

80. Circulation 62, III-80 (1980).

81. Circulation 62, III-76 (1980).

82. In Clinical Methods: The History, Physical and Laboratory Examinations (eds Walker, H. K., Hall, W. D. & Hurst, J. W.) 1122-1161 (Butterworth, Boston) (1980).

83. In The Heart, Update III (ed. Hurst, J. W.) 1-8 (McGraw-Hill, New York) (1980).

84. In Heart Disease: A Textbook of Cardiovascular Medicine, vol. 2 (ed. Braunwald, E.) 1517-1582 (Saunders, Philadelphia).

85. Primary Cardiol. 6, 101-109 (1980).

86. Primary Cardiol. 6, 112-118, 122 (1980).

87. New Engl. J. Med. 304, 129-135 (1981).

88. Ann. Surg. 194, 189-192 (1981).

89. Circulation Res. 49, 1017-1028 (1981).

90. Circulation 63, 29-35 (1981).

91. Am. J. Physiol. 240, H399-H407 (1981).

92. J. clin. Invest. 68, 225-239 (1981).

93. Archs Path. Lab. Med. 105, 403-406 (1981).

94. Am. J. Cardiol. 48, 702-710 (1981).

95. Am. J. Physiol. 241, H591-H599 (1981).

96. Proc. natn. Acad. Sci. USA 78, 7152-7156 (1981).

97. Clin. Res. 29, 444A (1981).

98. Clin. Res. 29, 428A (1981).

99. Chest 80, 388 (1981).

100. Circulation 64, IV-280 (1981).

101. Circulation 64, IV-210 (1981).

102. Circulation 64, IV-69 (1981).

103. Circulation 64, IV-68 (1981).

104. Fedn Proc. 40, 485 (1981).

105. Clin. Res. 29, 562A (1981).

106. Clin. Res. 29, 270A (1981).

107. Clin. Res. 29, 270A (1981).

108. Clin. Res. 29, 184A (1981).

109. Clin. Res. 29, 184A (1981).

110. Clin. Res. 29, 184A (1981).

111. Z. Kardiologie 70, 317 (1981).

112. J. nucl. Med. 22, P53 (1981).

113. Chest 80, 385-386 (1981).

114. Circulation 64, IV-280 (1981).

115. Circulation 64, IV-236 (1981).

116. Circulation 64, IV-222 (1981).

117. Circulation 64, IV-178 (1981).

118. Circulation 64, IV-122 (1981).

119. Circulation 64, IV-99 (1981).

120. In The Heart, Update IV (ed. Hurst, J. W.) 219-286 McGraw-Hill, New York) (1981).

121. Am. Heart J. 102, 807-808 (1981).

122. Am. Heart J. 101, 124-126 (1981).

123. J. cardiovasc. Med. 6, 705-715 (1981).

124. Acta med. scand., Suppl. 651, 123-132 (1981).

125. J. cardiovasc. Med. 6, 1173-1177, 1181 (1981).

126. Am. J. Cardiol. 49, 1023 (1982).

127. New Engl. J. Med. 308, 1419 (1982) (retraction).

128. Ann. intern. Med. 99, 275-276 (1982) (retraction).

129. Circulation Res. 53, 837 (1982) (retraction).

A#79. Heymsfield, S. B., and Glenn, J. F. (1983). Retraction. New Engl. J. Med. 308, 1400.

A#127. -. (4 April 1979). "Armored hearts and non-coronary infarctions: unexplained features of pancreatitis." Abstract of report presented in annual Lloyd Hyde Medical House Staff and Fellows Research Program, Emory University School of Medicine.

A#129. Tosteson, D. C. (25 January 1982). Letter to members of the Faculty of Medicine, Harvard Medical School, 1 page.

A#130. Ross, R. S., Barger, A. C., Benacerraf, B., Dreben, B. S., Farber, S. J., Frug, G., Levy, R. I., and Martin, J. B. (29 January 1982). Report of an ad hoc advisory committee to the dean of the Harvard Medical School on dishonesty in scientific research. Harvard University Gazette 77, 1, 11-12.

A#131. Hall, E. C., Huguley, C. M., Jr., Symbas, P. N., and Moran, N. C. (28 March 1983). "Report of ad hoc committee to evaluate research of Dr. John R. Darsee at Emory University." pp.1-32.

A#132. Moran, N. C. (5 May 1983). "Addendum to report on research of Dr. John R. Darsee at Emory University School of Medicine." pp. 1-3.

A#133. Morgan, H. E., Klocke, F. J., Shepherd, J. T., and Wildenthal, C. K. (29 June 1982). "Report of the National Heart, Lung, and Blood Institute special panel to review

alleged misconduct at Brigham and Women's Hospital/Harvard Medical School, Boston, Massachusetts." pp. 1-36 and 7 Appendices.

A#134. -. (21 September 1982). "Comments concerning NHLBI panel's report and range of possible NIH actions." Memorandum to NIH staff members, pp. 1-36 and 6 Appendices.

A#135. -. (14 September 1982). Memorandum to W. F. Raub, NIH. pp.1-21.

A#147. Sierles, F., Hendrickx, I., and Circle, S. (1980). Cheating in medical school. J. Med. Educ. 55, 124-125.

A#148. Zastrow, C. H. (1970). Cheating among college graduate students. J. Educ. Res. 64, 157-160.

A#149. Broad, W., and Wade, N. (1982). Betrayers of the Truth (New York: Simon and Schuster).

A#150. Zuckerman, H. (1977). Deviant behavior and social control in science. In Deviance and Social Change, E. Sagarin, ed. (Beverly Hills, California: Sage), pp. 87-138.

A#151. Cannon, W. B. (1945). The Way of an Investigator: A Scientist's Experiences in Medical Research (New York: Norton), p. 93.

A#152. Wolins, L. (1962). Responsibility for raw data. Am. Psychol. 17, 557-658.

A#153. Broad, W. J. (1980). Imbroglio at Yale (I): emergence of a fraud. Science 210, 38-41.

A#154. Wade, N. (28 February 1983). Hot dog science. New York Times, p. A14.

A#155. Emory University. (6 May 1983). "Emory Announces Committee Findings." News release. pp. 1-4.

A#156. Knox, R. (1983). The Harvard fraud case: where does the problem lie? J. Am. Med. Assoc. 249, 1797-1807.

A#157. Knox, R. A. (1983). Deeper problems for Darsee: Emory probe. J. Am. Med. Assoc. 249, 2867-2876.

A#158. Knox, R. A. (23 May 1983). Medical fraud: rise and fall of a medical legend. Boston Globe, pp. 41, 46.

A#159. Broad, W. J. (1982). Harvard delays in reporting fraud. Science 215, 478-482.

A#160. Broad, W. J. (1982). Report absolves Harvard in case of fakery. Science 215, 874-876.

A#161. Culliton, B. J. (1983). Coping with fraud: the Darsee case. Science 220, 31-35.

A#162. Culliton, B. J. (1983). Emory reports on Darsee's fraud. Science 220, 936.

A#165. Short, E. M. (1983). L'affair Darsee. Clin. Res. 31, 448-449.

A#166. Friesinger, G. C., Motulsky, A. G., and Brown, D. M. (14 April 1983). Report of External Review Committee to the Dean of Emory University School of Medicine.

A#167. Bernhoft, R. A., Pellegrini, C. A., and Way, L. W. (1982). Peritoneovenous shunt for refractory ascites. Arch. Surg. 117, 631-634.

A#168. Gullstrand, P., Alwmark, A., Brjesson, B., Joelsson, B., and Bengmark, S. (1982). Peritoneovenous shunting for intractable ascites. Scand. J. Gastroenterology 17, 1009-1012.

A#169. Epstein, M. (1982). Peritoneovenous shunt in the management of ascites and the hepatorenal syndrome. Gastroenterology 82, 790-799.

A#170. -, and -. (1982). J. Clin. Invest. 70, 917.

A#171. [-.] (1982). Proc. Nat. Acad. Sci. USA 79, 6390.

A#172. [Huth, E. J.] (1982). Authorship from the reader's side. Ann. Intern. Med. 97, 613-614.

A#173. Weinstein, D. (1979). Fraud in science. Social Sci. Quart. 59, 639-652.

A#174. Manwell, C., and Baker, C. M. A. (1981). Honesty in science: a partial test of a sociobiological model of the social structure of science. Search (Sydney) 12, 151-160.

A#175. Baird, J. S., Jr. (1980). Current trends in college cheating. Psychology in the Schools 17, 515-522.

A#176. Schab, F. (1980). Cheating among college and non-college bound pupils, 1969-1979. Clearing House 53, 379-380.

A#177. Bowers, W. J. (1964). Student Dishonesty and Its Control in College (New York: Bureau of Applied Social Research, Columbia University).

A#178. Harp, J., and Taietz, P. (1966). Academic integrity and social structure: a study of cheating among college students. Social Probl. 13, 365-373.

A#180. Cohn, V. (8 October 1983). Doctor test invalidated by cheating. Washington Post, pp. A1, A10.

A#181. Science Citation Index, Source Index. (1979-1984). (Philadelphia: Institute for Scientific Information).

A#182. Milnor, W. R. (1982). Hemodynamics (Baltimore: Williams and Wilkins), p. 17.

A#183. Glenn, J. F. (1983). Validity of publication. Clin. Res. 31, 451.

A#184. Swazey, J. P., and Scher, S. R. (1981). The whistleblower as a deviant professional: professional norms and responses to fraud in clinical research. In Whistleblowing in Biomedical Research, J. P. Swazey, and S. R. Scher, eds. (Washington: U. S. Government Printing Office), pp. 173-192.

A#185. Oakes, A. N. (1981). Protecting the rights of whistleblowers and the accused in federally supported biomedical research. In Whistleblowing in Biomedical Research,

J. P. Swazey, and S. R. Scher, eds. (Washington: U. S. Government Printing Office), pp. 111-142.

A#186. Westin, A. F. (10 January 1982). Faculty research and the whistle blowers. New York Times, Section 13, p. 61.

A#187. Edsall, J. T. (1981). Two aspects of scientific responsibility. Science 212, 11-14.

A#188. Raub, W. F. (1981). The National Institutes of Health's perspective on misconduct by grantees. In Whistleblowing in Biomedical Research, J. P. Swazey, and S. R. Scher, eds. (Washington: U. S. Government Printing Office), pp. 151-154.

A#189. Hull, D. L. (1978). Altruism in science: a sociobiological model of co-operative behaviour among scientists. Anim. Behav. 26, 685-697.

A#190. Merton, R. K. (1957). Social Theory and Social Structure (Glencoe, Illinois: Free Press), p. 559.

A#192. Sklar, J. (1982). Fraud in science. Clin. Res. 30, 435-436.

A#193. [Huth, E. J.] (1983). Responsibilities of coauthorship. Ann. Intern. Med. 99, 266-267.

A#194. Burman, K. D. (1982). "Hanging from the masthead": reflections on authorship. Ann. Intern. Med. 97, 602-605.

A#195. Gjerde, C. L., and Colombo, S. E. (1982). Promotion criteria: perceptions of faculty members and departmental chairmen. J. Med. Educ. 57, 157-162.

A#197. Weinstein, D. (1981). Commentary: bureaucratized science and resistance to whistleblowing. In Whistleblowing in Biomedical Research, J. P. Swazey, and S. R. Scher, eds. (Washington: U. S. Government Printing Office), pp. 193-196.

A#198. Angell, M. (1983). Fraud in science. Science 219, 1417-1418.

A#199. Gould, S. J. (1978). Morton's ranking of races by

cranial capacity. Science 200, 503-509.

A#200. Pemberton, C. (29 November 1983). Institutional research study 83-32: Results from the spring 1983 student and faculty surveys on academic honesty at the University of Delaware, pp. 1-32.

A#201. Mahoney, M. J. (1976). Scientist as Subject: The Psychological Imperative (Cambridge, Massachusetts: Ballinger), pp. 79-82.

A#202. "Information to Contributors." (January 1981). Proc. Nat. Acad. Sci. USA 78, ix.

A#205. St. James-Roberts, I. (25 November 1976). Cheating in science. New Scient. 72, 466-469.

A#206. Majerus, P. W. [Editor, Journal of Clinical Investigation]. (1981). Letter of 28 May 1980 to P. Felig. Reproduced in Fraud In Biomedical Research. Hearings before the Subcommittee on Investigations and Oversight of the Committee on Science and Technology, U. S. House of Representatives, Ninety-Seventh Congress. March 31-April 1, 1981 (Washington: U.S. Government Printing Office), p. 95.

A#207. Handler, P. (1981). Statement. In Fraud in Biomedical Research. Hearings before the Subcommittee on Investigations and Oversight of the Committee on Science and Technology, U. S. House of Representatives, Ninety-Seventh Congress. March 31-April 1, 1981 (Washington: U. S. Government Printing Office), pp. 10, 11.

A#208. Malone, T. E. (19 April 1983). "Fostering Integrity in Research: A Collective Responsibility." Presentation in Dallas, Texas, pp. 1-10.

A#209. Zuckerman, H., and Merton, R. K. (1971). Patterns of evaluation in science: institutionalization, structure and functions of the referee system. Minerva 9, 66-100.

A#210. Ingelfinger, F. J. (1974). Peer review in biomedical publication. Am. J. Med. 56, 686-692.

A#211. Peters, D. P., and Ceci, S. J. (1982). Peer-review practices of psychological journals: the fate of published articles, submitted again. Behavioral and Brain Sciences 5, 187-195, 252-255; commentary by other authors, 196-252.

A#227. Last, J. M., ed. (1983). A Dictionary of Epidemiology (New York: Oxford University Press), p. 23, 94.

A#228. Brandt, E. N., Jr. (1983). PHS perspectives on misconduct in science. Publ. Health Rep. 98, 136-140.

A#229. Lane, B. G., and Lipmann, F. (1961). Nonparticipation of ribonucleic acid in glutathione and ophthalmic acid synthesis. J. Biol. Chem. 236, PC80-PC81.

A#230. Simpson, M. V., Skinner, D. M., and Lucas, J. M. (1961). On the biosynthesis of cytochrome c. J. Biol. Chem. 236, PC81-PC82.

A#231. Hixson, J. (1976). The Patchwork Mouse (New York: Doubleday), pp. 146-148.

A#232. Racker, E. (1981). Warburg effect revisited. Science 213, 1313.

A#234. Woolf, P. (October 1981). Fraud in science: how much, how serious? Hastings Center Rep. 11, 9-14.

A#235. Jackson, C. I., and Prados, J. W. (1983). Honor in science. Am. Scient. 71, 462-464.

A#236. Greenberg, D. A. (1984). Scientific fraud. Am. Scient. 72, 118.

A#237. Van Valen, L. (1976). Dishonesty and grants. Nature 261, 2.

A#238. Gratzer, W. (1983). Trouble at t'lab. Nature 302, 774-775.

A#239. Relman, A. S. (1983). Lessons from the Darsee affair. New Engl. J. Med. 308, 1415-1417.

A#240. Roth, J. A. (1966). Hired hand research. Am. Sociol. 1, 190-196.

A#241. Croll, R. P. (1984). The noncontributing author: an issue of credit and responsibility. Perspect. Biol. Med. 27, 401-407.

A#242. Stetten D. (1984). Reported laboratory frauds in biomedical science. Science 226, 1374, 1376.

A#243. Krevans, J. R., et al. (1982). The maintenance of high ethical standards in the conduct of research. J. Med. Educ. 57, 895-902.

A#244. Bailar, J. C., III, and Patterson, K. (1985). Journal peer review: the need for a research agenda. New Engl. J. Med. 312, 654-657.

A#245. Pigman, W., and Carmichael, E. B. (1950). An ethical code for scientists. Science 111, 643-647.

A#246. Clapp, J. (1974). Professional Ethics and Insignia (Metuchen, New Jersey: Scarecrow Press).

A#247. "Yale Policy Statement on Collaborative Research." (13-20 September 1982). Yale Weekly Bulletin and Calendar 11, 9.

A#248. Hastings Center Staff, eds. (1984). The Hastings Center's Bibliography of Ethics, Biomedicine, and Professional Responsibility (Frederick, Maryland: University Publications of America), pp. 10-11.

A#249. Danforth, W. H., et al. ([April 1983]). Report of the Association of American Universities Committee on Integrity of Research (Washington, D. C.: Association of American Universities), pp. 1-6.

A#250. "The Chemist's Creed." (14 September 1965). American Chemical Society.

A#251. Majerus, P. W. (1982). Fraud in medical research. J. Clin. Invest. 70, 213-217.

A#252. Sigma Xi. (1984). Honor in Science (New Haven, Connecticut: Sigma Xi, The Scientific Research Society), pp. 1-38.

A#253. Chalk, R., Frankel, M. S., and Chafer, S. B. (1980). Professional Ethics in the Scientific and Engineering Societies (Washington, D. C.: American Association for the Advancement of Science), pp. 1-224.

A#254. Stossel, T. P. (1985). Speed: an essay on biomedical communication. New Engl. J. Med. 313, 123-126.

A#255. Petersdorf, R. G. (1986). The pathogenesis of fraud in medical science. Ann. Intern. Med. 104, 252-254.

A#256. Woolf, P. K. (1986). Pressure to publish and fraud in science. Ann. Intern. Med. 104, 254-256.

A#257. Huth, E. J. (1986). Irresponsible authorship and wasteful publication. Ann. Intern. Med. 104, 257-259.

A#258. Bailar, J. C., III. (1986). Science, statistics, and deception. Ann. Intern. Med. 104, 259-260.

A#259. Angell, M. (1986). Publish or perish: a proposal. Ann. Intern. Med. 104, 261-262.

A#260. Huth, E. J. (1986). Abuses and uses of authorship. Ann. Intern. Med. 104, 266-267.

A#261. Huth, E. J. (1986). Guidelines on authorship of medical papers. Ann. Intern. Med. 104, 269-274.

A#262. Miers, M. L. (1985). Current NIH perspectives on misconduct in science. Am. Psychologist 40, 831-835.

A#263. Balzer, H. D. (1985). Is less more? Soviet science in the Gorbachev era. Issues in Science and Technol. 1, 29-46.

Table 1. Errors and discrepancies in published papers

Ref	Source	Errors	Discrepancies
6	E	0	1(17)
24	E	3	1(3)
25	E	0	0
26	E	4	0
50	E	3	5(13), 2(98)
51	E	14	0
52	E	13	1(9), 2(15)
87	E	8	8(19), 8(28), 4(40)
88	E	8	1(20), 1(30)
89	E	15	2(46), 2(55)
53	H	1	0
90	H	13	2(67)
91	H	6	4(74), 4(108)
92	H	37	1(70), 1(78)
93	H	0	0
94	H	10	4(76)
95	H	11	4(96), 3(105)
96	H	7	4(95), 3(105)

Abbreviations and Description:
Ref (column 1): Reference number
Source (col. 2): E=Emory, H=Harvard
Errors (col. 3): Number of errors in reference
Discrepancies (col. 4): Number of discrepancies with indicated reference
Eighteen research papers are listed individually by their reference numbers. **Example** (first line of table): Ref. 6, a research paper from Emory, was found by us to contain no errors and to contain 1 discrepancy (with ref. 17).

Table 2. Discrepancies between a paper[88] and two abstracts[20,30]

	Ref. 88	Ref. 30	Refs. 20 and 88	Refs. 20 and 30
Mean or s.d.	m.±s.d.	m.±s.d.	s.d.	s.d.
Parameter measured:				
CO	2.1±0.8	2.6*±0.8	1.0**	0.9
PCWP	10.0±3.7	12.2*±2*	4.4	1.8
PVR	-43±33	-64*±24	30	43***
SVR	-123±33	-192*±168*	199**	158
SWI	10.6±4.2	11.3±5.8*	4.8**	5.2***

Discrepancies are indicated between columns 2 and 1 (*); columns 3 and 1 (**), and columns 4 and 2 (***). Values in col. 1 were obtained by measurements of the 0.5 hr points in Figure 1 of ref. 88. The values in col. 2 were copied directly from ref. 30, except that for SVR and PVR, the R units[30] have been converted to dyne s cm-5 (factor of 80; ref. A#182). Columns 3 and 4 contain standard deviations for 5 patients calculated from the means and standard deviations given in ref. 20 for 4 patients and the means given in refs. 88 or 30 for 5 patients. The values were derived from calculations based on "population" standard deviations; calculations based on "sample" standard deviation gave similar values and conclusions. When comparing values, allowance was made for rounding of the final digit in refs. 20 and 30 and for uncertainties of measurement in ref. 88.
Abbreviations: CO, cardiac output; PCWP, pulmonary capillary wedge pressure; PVR, pulmonary vascular resisance; SVR, systemic vascular resistance; SWI, stroke work index. Units are given in the publications, except as noted above.

Table 3. Authors: number of publications and practices

Publications	Auth	Total: A	Total: B	Practice
Total = 1-2				
	1	0	0	
	2	0	0	
	3	0	0	
	4	0	0	
	5	0	0	
	6	0	0	
	7	0	0	
	8	0	0	
	9	0	0	
	10	0	0	
	11	1	0	3(1)
	12	1	0	3(1)
	13	1	0	1(1)
	14	1	0	1(1)
	15	1	0	1(1)
	16	1	0	1(1)
	17	1	0	1(1)
	18	1	0	1(1)
	19	1	0	1(1)
	20	0	1	8(1)
	21	0	1	9(1)
	22	1	1	1(1),5(1)
	23	4	0	1(1),2(3)
Total = 3-5				
	24	0	0	
	25	0	0	
	26	1	0	2(1)
	27	1	0	2(1)
	28	1	0	2(1)
	29	0	1	10(1)
	30	0	1	7(1)
	31	4	0	1(1),2(3)

32	4	0	1(1),2(3)
33	4	0	1(1),2(3)
34	4	0	1(1),2(3)
35	4	0	1(1),2(1)[2],3(1),4(1)
36	5	0	1(2),2(2)[2],3(1)
37	4	1	1(3),2(1),5(1)

Total = 6-15

38	7	0	1(2),2(4),3(1)
39	6	1	1(2),2(3),3(1),7(1)
40	7	1	1(2),2(3),4(2),7(1)
41	6	4	1(3),2(3),5(2),6(1),7(1)
42	6	5	1(3),2(3),5(2),6(1),7(2)

Total = 16-41

43	11	1	1(4)[1],2(5)[2],3(1),4(1),7(1)
44	23	0	1(9)[1].2(7)[2],3(4),4(3)
45	24	0	1(8)[3],2(9)[2],3(3),4(4)
46	21	9	1(7),2(9),4(5),5(5),6(1),7(2),10(1)
47	23	12	1(13),2(10),5(7),6(1),7(3),10(1)

Abbreviations and Description:

Publications (column 1): Range containing the total number of each author's publications (papers, abstracts, chapters)

Auth (col. 2): Author's numerical designation

Total (cols. 3 and 4): Total type A and type B practices (see text)

Practice (col. 5): Practice number (see Table 6), in bold face, followed by total number of instances of that practice, in parentheses, (). Brackets, [], indicate that the author has stated (A#131) that he was unaware of the existence of abstracts before they were published (see **Appendix**); bracketed figures were not included in the totals (cols. 3 and 4).

Example. Author 38 was a coauthor of publications whose total number was between 6 and 15. There were 7 instances of a type A practice; there were none of type B. There were 2 instances of practice #1, 4 of #2, and 1 of #3.

Table 4. Frequency of practices classified by publication and by coauthor

Type of practice	Number of publications	Number of coauthors
1	28	26
2	42	21
3	7	9
4	13	6
5	7	6
6	1	4
7	8	8
8	1	1
9	1	1
10	2	3

Column 1: Type of practice as defined in Table 6.
Col. 2: Number of publications by coauthors in which there were one or more examples of the practice in col. 1.
Col. 3: Number of coauthors involved in each practice. Some publications and some coauthors were associated with more than a single type of practice.

Table 5. Number of dogs for which results were shared

	Ref. 53	Ref. 90	Ref. 91	Ref. 92	Ref. 94
Ref. 53	33	8	12	0	9
Ref. 90	8	22	9	0	11
Ref. 91	12	9	76	6	7
Ref. 92	0	0	6	48	0
Ref. 94	9	11	7	0	44

The boldface figures (on diagonal) give the total numbers of dogs used in each study. The remaining entries give the numbers of dogs for which experimental data were shared between pairs of studies. Although not shown in this table, data for the same 2 dogs were used in 4 studies[53,90,91,94]. Reference A#133 was the source of this table, which in turn was based on information provided by one of the senior coauthors at Harvard.

Table 6. Practices departing from generally accepted standards

Type A.

1. An author fails to correct errors in a publication. We counted as errors only those that altered the factual content of a paper (usually numerical values).

2. An author fails to check a publication for consistency with the research group's previously published data.

3. An author fails to retain names or identifying numbers for human subjects included in a published paper.

4. An author accepts an authorship when he had little or no actual participation in the acquisition and interpretation of data and did not make an intellectual contribution which was essential to the research -- honorary coauthorship (see text).

Type B.

5. An author publishes a statement that he knows or should know is inaccurate or misleading.

6. A scientist coauthors a paper which misleadingly cites a previous paper containing data being republished, in such a way that the true relationship between the two papers is obscured.

7. An author publishes very similar abstracts under very different titles; as a result one who reads only the titles may be led to believe the abstracts are different.

8. A scientist has unique knowledge that makes possible the recognition of certain serious errors in someone's published work -- errors not discoverable by anyone else. The scientist who discovers these errors refrains from informing the appropriate person about them.

9. A scientist publishes a substantial amount of information which he has received from someone else, information whose source he does not acknowledge in print.

10. A senior scientist with substantial administrative responsibilities receives a complaint that a junior scientist in his department may be involved in questionable data collection. The problem eventually proves not to have been resolved satisfactorily. Meanwhile the senior scientist and the junior scientist coauthor a publication.

Table 7. Differences between experimental procedures in five papers

	Ref. 53	Ref. 90	Ref. 92	Ref. 94	Ref. 91
1. Coronary artery occlusion (time)	2-6 hr	6 hr	1 or 3 hr	4-6 hr	1 min-2 hr
2. Randomization before/after coronary artery occlusion	After	After	Before	After	Bef.(?)
3. Time from removal of coronary artery clamp to end	~1 min-4 hr	~1 min	6 or 75 hr	~1 min-2 hr	6 hr
4. Antiarrhythmic agent:	Lido	Lido	Lido	Proc'de	None
a. Initial bolus inj., mg/kg	1.5	1.5	1.5	3	None
b. Further intermittent inj.	Yes	No(?)	No(?)	No(?)	No
c. Continuous infusion	No	No	No	Yes	No
5. Dye administered initially	Fln	MeB,GV	Fln,GV	Fln	Fln,GV
6. Dye injected terminally	TS,MeB	MeB	MeB	MonB	MonB
7. Drug or NaCl soln., ml/kg	0(?)	1	1	2	0(?)
8. Catheter in aorta/left carotid	Carotid	Carotid	Carotid	Aorta	Carotid
9. A set of multiple procedures, (a)-(e), listed below.	No	No	Yes	No	Yes

There are substantial differences in experimental procedures for at least nine of the ten pairs of papers. The procedures in item 9 include: (a) insertion of micromanometer through stab wound in ventricular apex, (b) application of coronary cuff, (c) repeated brief reocclusion of coronary artery, (d) implantation of crystal transducers in myocardium, and (e) delay in examination of heart after death (for flowmeter calibration). Abbreviations: Lido, Lidocaine; Proc'de, Procainamide; Fln, Fluorescein (injection); MeB, Methylene Blue (injection); GV, Gentian Violet (topical); TS, Thioflavine S; MonB, Monastral Blue. Information in the table comes from the papers themselves[53,90-92,94]; see also ref. A#133, appendix 5.

Table 8. Opening paragraph of reference 96

Previous studies from this laboratory have shown that brief (15 min) coronary occlusions followed by reperfusion do not cause necrosis but do result in prolonged abnormalities in myocardial biochemistry, function, and ultrastructure (1-3). At 72 hr after release of a 15-min coronary occlusion, the ATP concentration in reperfused previously ischemic myocardium was significantly reduced (by 22% of normal), the percentage of systolic shortening of reperfused left ventricular segments was reduced (by 42% of normal), and ultrastructural abnormalities were present, despite the absence of necrosis (1, 3). However, it was not known whether these changes ever were reversed to normal after longer periods of reperfusion and, if so, what the time course of the recovery is. Recently there has been renewed interest in the treatment of acute myocardial infarction by means of coronary reperfusion, either by injecting fibrinolytic agents directly into the obstructed coronary artery or by coronary revascularization (4, 5).

This is the opening paragraph, printed here verbatim, of a research paper[96] discussed in the text of our report. Reference 3 in the paragraph above is the only reference in this paper[96] to the authors' own previous (72 hour) study[95], in which some of the same animals were used and some of the same results were reported.

Figure Legends

Figure 1. Pedigree Reproduced from Reference 87

The age of each family member at the start of the investigation is given below each symbol. Arrows (second row) denote those who had died before the start of the investigation. Note the 17-year-old male (row 4) with children ages 8, 7, 5, and 4; at the age of 8 or 9 (probably 8) he impregnated the mother of his oldest child.

Figure 2. Reproduction of Figure 1 from Reference 88

The graph depicts the mean and standard deviation for changes in haemodynamic values of all five patients in the study. Note the remarkable similarity in size of many error bars. Note also the relative constancy of SVR (average value for five patients) at -160 for four serial measurements, despite the variability of SVR (individual values for two patients) in Figures 2 and 3 of ref. 88 (not reproduced here).

79

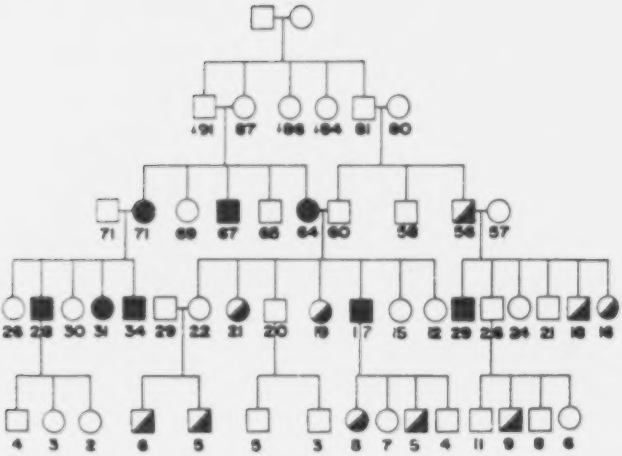


Fig. 1.
Stewart and Feder

80

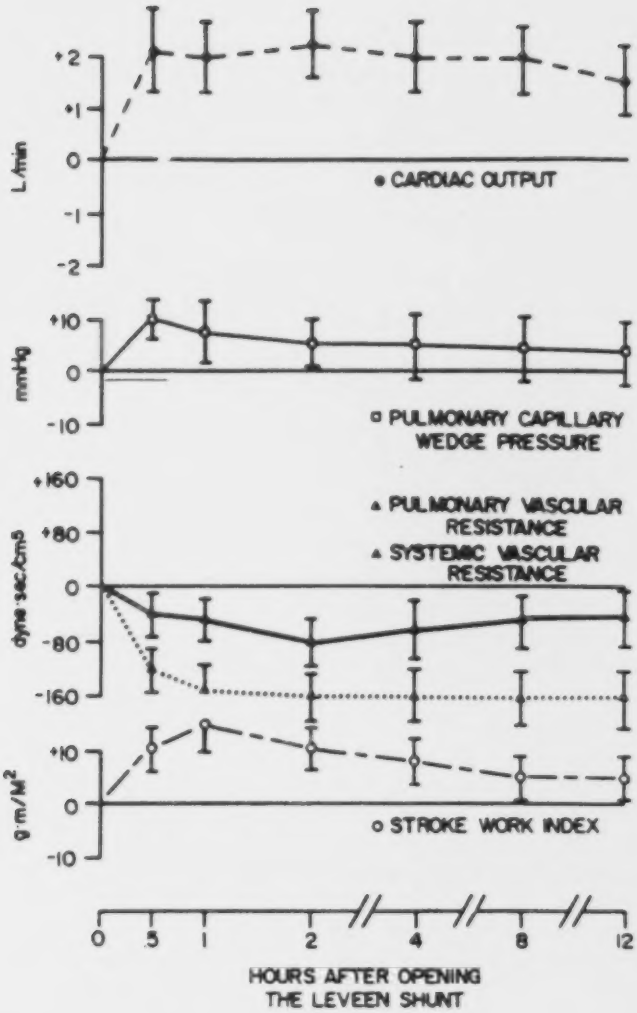


Fig. 2.
Stewart and Feder

DISCUSSION

Mr. FUQUA. Thank you, Mr. Stewart and Dr. Feder.

Is this just limited, the malpractice, to the field of biomedical research or do you think it is found in most of the other disciplines?

Mr. STEWART. Perhaps I could answer that question by saying the only thing we know for certain, within the accuracy of our study, is what happened with these 47 people. The only reason we studied them was because it was a sample—then you get into the question of how good a sample and a sample of what? We describe it as a sample of biomedical scientists because they were biomedical scientists. But I think it is certainly possible that some of these same problems are not unique to biomedical science but may possibly be present in other forms of science as well.

Mr. FUQUA. Did you interview any of the 47 people whose reports you studied?

Mr. STEWART. Yes, we did. We talked to as many of those people as we could.

Mr. FUQUA. How many would that be?

Mr. STEWART. That would be approximately 15. That is my guess. That is just an "off the top of my head" guess.

Mr. FUQUA. Did the others refuse to talk with you or were unavailable?

Mr. STEWART. A problem emerged in the sense that they felt that our talking to members of one university posed a problem and we were asked not to talk to anyone further. But we were not able to talk to everybody.

Mr. FUQUA. That was just one, though.

Mr. STEWART. One whole university. Since there were only two, one was deemed to be a large number. We, therefore, were not able to complete discussions with everybody. We relied on published literature, but we realized that the published literature might have errors in it and we wanted to talk to people to eliminate that possibility.

Mr. FUQUA. Was there a timeframe of the study that you conducted? I am talking about in the report. My question is really, is this something that you think is new or has it been going on for a long period of time in science?

Mr. STEWART. I don't know if we could answer that.

Ned, do you have a feeling?

Mr. FUQUA. What was the time span of the published reports that you reviewed? Were they recent, or did some go back several years?

Mr. STEWART. Some 3 or 4 years would be my impression.

Ned, is that about right?

Dr. FEDER. The papers that we examined were published in, I think, 1978 to 1981.

Mr. FUQUA. So it is about a 3-year timeframe?

Dr. FEDER. That is right. And we chose that particular timeframe because our sample forced us to. Those were the years that these particular people published with the person we were studying regarding the forging of data.

Mr. FUQUA. But do you think this existed prior to 1978? Ten years ago? 1968?

Dr. FEDER. Our study really doesn't deal with that question.

Mr. FUQUA. Well, what do you hear as you talk among your science colleagues?

Dr. FEDER. Well, our best evidence for the generality of our study does come from the comments of colleagues who have read it. And the comments are generally not those of astonishment or surprise. They may say things, most of them, such as I am not surprised or I didn't realize it was that high. But very few of them say you are way off.

Mr. STEWART. Mr. Chairman, I will stick my neck out on that if I can. Though we don't have any precise data, it is our clear impression that these problems didn't originate all of a sudden in 1978, so I think they have been going on for a long time.

Mr. FUQUA. Do you have any indication that this is true in other countries, rather than just the United States?

Mr. STEWART. That one I am not going to stick my neck out on. Our data don't bear on it and our experience isn't adequate to answer that.

Mr. FUQUA. What about self-imposed in-house review procedures to make sure that these types of alleged misconduct do not occur? Is there any in-house review of papers and so forth?

Mr. STEWART. Again, I am not an expert on that subject and can't give you a complete answer. Certainly where we work there is a procedure, but I don't believe anyone believes that it would be adequate to pick up this sort of misconduct. I think mainly the procedure whereby two people read the paper and signoff on it is designed to prevent release of material that might be sensitive in some way, or grossly incompetent. I think that adequate in-house review just isn't in place at the moment to deal with this and might or might not be the best way to deal with it.

Does that answer the question?

Mr. FUQUA. What about some of the professional societies that these people may belong to, do they do any policing—I don't mean policing, but self-discipline?

Mr. STEWART. I think that the normal mechanisms for reviewing publications don't include the medical societies. It passes, for instance, through some sort of an informal review within a department of, let's say, a scientists' immediate colleagues, and then it will go typically to a journal where it is again reviewed. I suppose societies become involved when it becomes clear that there is a problem, and then they may look into it. But I don't think they customarily review publications.

Mr. FUQUA. You indicated that based on your interview with some of the senior scientists that some of the misconduct may be quite widespread and that the institutional mechanisms for dealing with the problems are inadequate. What type of institutional mechanisms do you suggest might be effective in trying to deal with this problem?

Mr. STEWART. Can I answer that in two separate pieces?

As we asked the senior scientists to review our paper, though we weren't aware of it, we were doing in essence a study within a study. We would call the scientists, many of whom were known to us only by reputation, introduce ourselves and say what it was we were interested in taking their time to do; namely, to read our

paper. In the course of making all the phone calls and contacting these people a surprising number of them would just talk to us and say, well, that is a fascinating subject because—and before we had realized it we had accumulated an enormous amount of data. Not of a formal scientific sort but of a very convincing anecdotal sort where people would say, "This has been going on for years. It is a tremendous problem. I don't know what to do about it." Certain senior scientists said they were even afraid to talk about it with their own colleagues. I was surprised that they would talk about it with us, but then again we were studying the problem and they may have felt more at ease in a scientific setting like that. So in that sense we feel that we do have, in addition to our actual data, which don't bear on that question, we feel we have accumulated some experimental information.

To answer the actual question, I don't really know. I mean, we do suggest certain things such as the enormous pressure to produce publications of any sort of quality, regardless of whatever quality or usefulness. We think that is a contributing factor.

Mr. FUQUA. Publish or perish.

Mr. STEWART. Exactly. And it does appear to us it is getting worse.

Mr. FUQUA. Are we putting too much emphasis on that?

Mr. STEWART. Oh, in my opinion, absolutely.

Mr. FUQUA. It is pushing the system to respond in that fashion?

Mr. STEWART. We are almost requiring people to be dishonest in certain situations. We now have scientists who have 500, 700, even 1,000 publications attributed to them, and it is not possible in my opinion to do this without being an honorary author in many cases. We are also saying that it is not necessary for scientists to do really good things to publish 1,000 papers, either. We believe that the misattribution of scientific credit, that is, granting it where there really is no reason to grant it, either because the work wasn't done at all or because the person getting the credit didn't do the work, we think that may be a pretty significant factor.

Mr. FUQUA. What does this do to the peer review process, then? We have had witnesses testify that in the peer review process that they place a great deal of weight, emphasis, on the track record of the individual that is submitting the proposal. And when you inquire further, well, it is publications and so forth, of what they have done in the past. Now would not the peer review have a chance to expose some of this fraud that may exist?

Mr. STEWART. I have a suggestion, and maybe Ned does. A possible question would be did you read the papers? You probably assumed they did and they may have.

Mr. FUQUA. Well, the reviewers say they barely have time to read the proposal.

Mr. STEWART. That is it. In fact, sometimes bibliographies are so long they can't be read. So "papers never get read" is an exaggeration. That isn't true. But it is certainly true they often don't get read.

One of the most shocking things in our whole study, that we didn't put in it because it wasn't in a certain sense subject to quantitative measurements, was the role of the scientist who was responsible for promoting the person who forged all the data, many

times over, having hired him and championed his work. When he found out that the researcher was a forger he then sat down and read his papers, and it was perfectly clear they were all forged. That coincides with our experience. Because when you look at these papers, they are obviously wrong papers. But the question that naturally comes up is why wasn't anyone reading these papers in the first place if it is so important to generate such large numbers of papers? We think that the emphasis on producing papers that are written but not read is not helpful and is, in fact, harmful.

Mr. FUQUA. But getting a paper published, doesn't somebody read the paper?

Mr. STEWART. You wonder. You don't always know. I mean, when we looked at these papers they were just full of errors. It is not at all clear that anyone reads them. It is certainly the case that many papers, I believe 50 percent roughly—Dr. Patricia Woolf may be able to correct me on this—some 50 percent of papers written are never cited. In the sample that we looked at it wasn't really clear that anybody was reading them, and they were published in very leading journals.

So there is a problem. So much is written nobody could read it all. I don't think it would be possible to read, I mean——

Mr. FUQUA. Like reading the Congressional Record.

Mr. STEWART. That's it. You got it.

Mr. FUQUA. Mr. Lewis.

Mr. LEWIS. Well, Mr. Chairman, you asked most of the questions that I had an interest in.

I would like to inquire as to after your research how many libel suits have you encountered or anticipate encountering?

Mr. STEWART. Even one would be too many to look forward to. We have not been sued at all. But then again we haven't published anything. We have 150 pages of lawyers' letters directed at us, at the journals, and at the NIH—but mainly at us—and we have had to answer every single one of those in detail. So we haven't been sued yet, but that can't be too much worse than answering all those letters. They are certainly indicating they will sue us.

Mr. LEWIS. I would think so.

Are you anticipating doing any followup on your study as to the damages possibly contained within the research community because of the erroneous material publicized?

Mr. STEWART. Our goal initially was to just do this study and get right back to our work—say to our colleagues this is an important subject, you can study it, it is a lot of fun, you know, come on in the water is fine. I am not sure that we have made that point convincingly because everybody says, "If I am going to get sued and if it is going to take such a long time and it can't get published, why should I do it?" We still are very interested in getting back to our own work, so that the last thing in the world we would like to do personally is to do any more of this. Once was more than enough.

Mr. LEWIS. Well, I think it would be interesting to find with erroneous papers and information and data contained within the papers that someone would accept this as gospel, move down with further research into an area and spend countless hours and up to years in this area.

Mr. STEWART. I hope Dr. Woolf will address that because we are not experts in this area. We did the study because we thought it was important and we thought we had a good idea. But I agree with you. I think it is a subject of enormous interest and importance. It is not being studied at all by anybody, almost. I would hope very much people would pick this up and do something with it. But because of our track record, I am not sure they will.

Mr. LEWIS. And in addition, the countless millions of dollars that will be awarded in grants for additional research that would be fallacious or misleading.

Mr. STEWART. That is quite true.

Mr. LEWIS. Thank you. Thank you, Mr. Chairman.

Mr. FUQUA. Thank you, Mr. Lewis.

I just have one final question. The two universities that you mentioned that had appointed I believe superficial and somewhat secret in nature committees to review these allegations, do you know if the Federal agencies involved in funding those projects, what they have done about trying to correct the situation? Has there been any followthrough?

Mr. STEWART. That is a bright side to the question. The two universities were Emory and Harvard, and it is my belief—and Ned can correct me if I am wrong—that there was little Federal money in Emory and a great deal of Federal money at Harvard.

The two university investigatory panels, as you indicated, had very serious deficiencies. The NIH formed its own panel and went in and looked at the Harvard case. And they were responsible for providing many of the facts on which we relied and they did a very careful, complete, thorough study.

Mr. FUQUA. So there was a followthrough in the case where the substantial Federal funds were involved?

Mr. STEWART. That is right.

Mr. FUQUA. Do you know of any other cases where there has been this type of followthrough when there has been allegations of misconduct?

Mr. STEWART. I can't answer that for certain. I am not aware of other cases. But what did emerge from our talking to people is many, many, many cases in which the evidence is covered up before it is brought to the attention of anybody. So it is our impression that it is extremely rare to have a study of this sort of thing happen at all.

Mr. FUQUA. Well, I am speaking of agencies such as the NSF or NASA or Department of Energy, or some of the other granting agencies.

Mr. STEWART. I don't know about other agencies, either.

Mr. FUQUA. Do you know if it is widespread among those agencies?

Mr. STEWART. I just don't know.

Mr. FUQUA. Your study was conducted primarily on this biomedical.

Mr. STEWART. Scientists. That is right. Scientists in general, and biomedical research scientists in particular. But as far as other agencies and other fields, I wouldn't know.

Mr. FUQUA. Gentlemen, thank you very much for sharing with us your investigation and your comments on that.

Mr. FUQUA. Our next witness will be Dr. Patricia K. Woolf, lecturer, Department of Sociology at Princeton University.

Dr. Woolf, I was called yesterday by a very mutual friend of ours, former Congressman Rogers. He said you would be a very creditable witness.

STATEMENT OF DR. PATRICIA K. WOOLF, LECTURER, DEPARTMENT OF SOCIOLOGY, PRINCETON UNIVERSITY, PRINCETON, NJ

Dr. WOOLF. Thank you, Mr. Fuqua.

I am Patricia Woolf, at present a lecturer in the Sociology Department of Princeton University. I have written and published on communication practices among scientists. For the last several years I have been studying scientific fraud and the role of journal editors in quality control in scientific research and publication. I was invited to participate in this hearing to discuss whether a manuscript describing a research project conducted by two NIH scientists, Ned Feder and Walter Stewart, contains evidence of serious problems in the way scientific research is currently conducted and communicated in America.

A brief answer is that their account is an original, thoughtful, and serious contribution to an important discussion. If it is accurate and representative, it clearly suggests that some reforms of research and publication practices are needed. Although serious questions have been raised about both the style and the substance of the manuscript, I believe that the processes of peer review, if allowed to function normally, were and are adequate to assess the potential value of the manuscript. Furthermore, science and the public would have been better served if these processes had been allowed to evaluate their research without the distorting effects of potential libel action.

Progress in science depends on maintaining attitudes, skepticism, and collegiality, which are not easily reconciled one with the other. I believe there is no other profession that makes such high demands on its practitioners in terms of the need for constant and critical skepticism about ideas and trustful cooperation of the colleagues who produce those ideas. Without collegiality and trust, the complex cooperative enterprise that characterizes modern science would be slowed to a standstill. There is little doubt that collegiality can make objective criticism extremely difficult and undermine the ability to maintain healthy skepticism. But skepticism is crucial to critical evaluation of the ideas, data and conclusions of scientific review. In ordinary life, skepticism undermines friendships; trust engenders friendships. Still, science demands that we dissociate our judgments about people from our acceptance of their ideas.

Collegiality depends on the assumption of integrity. A serious question for science is whether this assumption interferes with our ability to be appropriately critical of our colleagues' research, and to be sufficiently diligent in detecting problems. A more serious question is how we can insure that the processes of scientific review and appraisal are sufficiently rigorous to minimize and correct errors and to eliminate fraud to the greatest possible extent.

In the past, I believe incidents of fraud or falsification in science were rare. Among the most notable were Piltdown man and Cyril Burt's studies of nonexistent twins. The painted mouse incident shocked the biomedical community in 1975. Then, in the summer of 1980, there were four cases of fraud which were disclosed and received widespread national publicity. Since then the scientific community has set itself the task of examining and improving its quality assurance programs. It has been prompted to do so by congressional hearings and by formal inquiries by the Federal agencies that fund research. But more than those reasons, it was prompted to do so by a profound sense of outrage and embarrassment on the part of scientists. Fraud in science had the unfortunate effect of undermining trust among scientists—among some scientists—but I believe it has had a compensating virtue of stimulating self-criticism. This has taken place locally at institutions where fraud has occurred and across the Nation.

It has become clear that responsible authorship of scientific papers is central to quality control in research and publication. The criteria for authorship have been appraised and definitions suggested by spokesmen representing the two major institutions of science: research universities and scientific journals. President Donald Kennedy's message, "On Academic Authorship," to the Stanford University faculty received attention in the scientific press and was distributed to members of the Association of American Universities. Scientific editors, among them Edward Huth of the *Annals of Internal Medicine* and Arnold Relman of the *New England Journal of Medicine* have taken the lead in America in specifying contributions that scientists should have made in order to assume authorship, and the credit and the responsibilities that authorship entails. The American Chemical Society has recently published a revised version of its code of ethics which includes criteria for authorship. The Council of Biology Editors and the American Psychological Association have written authorship policies. Unanimous agreement among the International Committee of Medical Journal Editors has produced a revision to their uniform requirements, a style and substance guide to which 300 journals currently adhere.

There is evidence, but it is anecdotal and not systematic, that there is now increased vigilance in collaborative efforts, in the preservation of data and in the reviewing and editing of manuscripts. One of the coauthors in a well-known case of fraud said that people were "tending to be more careful," especially in multiinstitutional collaborative studies. New precautions are being taken with respect to data storage, data analysis and systematic recordkeeping.

In addition, researchers are trying to be more scrupulous in checking the credentials of potential research collaborators. Although it is not a congenial attitude, I expect that in view of the recent cases of fabrication of data, coworkers and reviewers will consider the possibility of fraud along with simple error as an explanation for discrepancies in scientific data.

It would be important to document such changes in attitude and practice among research scientists. But systematic empirical studies of scientists' research and communications behavior are ex-

tremely rare. In the conduct of such studies it will be important to avoid the imputation that because research practices in a given laboratory have been tightened up that the prior behavior was corrupt.

Most of the self-corrective measures taken by scientists have taken the form of examinations of conscience and re-examinations of standard practices. These appear to be based on the belief that individual integrity and collegial vigilance are the most efficient means of assuring quality and honesty in science. I think that is the case. There have been several symposia sponsored by the AAAS, such as the recent one on "Pressure to Publish and Fraud" and a continuing project on "Openness and Secrecy in Science and Technology." There have been sessions of the Society for Scholarly Publishing and the Council of Biology Editors devoted to the prevention of fraud in science and ways to mitigate its effects. Scholarly research on these topics has been funded by the National Science Foundation.

The manuscript by Stewart and Feder, "Professional Practices Among Biomedical Scientists: A Study of a Sample Generated by an Unusual Event," has been called the most famous unpublished paper in America. It is one of many responses to an egregious instance of fraud in science. Like other studies it has gone beyond criticism of flagrant plagiarism and falsification of data to include other misconduct in science such as duplicate publication, breaking the report of one experiment into several shorter manuscripts, failure to maintain careful records and raw data, and using statistical or graphical techniques to make poor data look better. These abuses appear to be for the purpose of publishing more journal articles, rather than for the advancement of scientific knowledge. They suggest that in some cases publication has become a mere device for professional advancement rather than a means of communicating new knowledge.

The Feder-Stewart paper is unlike other studies in one sense. It is unique in that it uses an analysis of published journal articles and reports of investigating committees to examine some of the research and publishing practices of the coauthors of those papers. It is important to note at the outset, as Feder and Stewart do, that the group of authors is not a random sample and may not be regarded as representative of any larger group of scientists. Furthermore, the Feder-Stewart paper does not discuss all of the publications of the scientists involved; it discusses only that research which was conducted and published in connection with a scientist who was later found to have forged scientific data.

Does the Feder-Stewart paper indicate that there is something seriously wrong in American research? First let me say that I believe their paper to be a conscientious attempt to answer a difficult question that most scientists and policymakers are interested in and most sociologists of science have found so difficult that they haven't even tried to answer it. That question is "How prevalent is misconduct in scientific research and publication?" When asked whether disclosed cases of fraud are the tip of the iceberg, sociologists shrug and reply that statistics on all kinds of deviance are notably hard to come by. The Feder-Stewart paper is a serious critique of American science.

If it is accurate, it should be a matter of grave concern to scientists themselves and to Federal agencies that fund science. I am not a cardiologist and have not conducted any bench research for more than 20 years, and therefore cannot adequately judge the technical details of some of the specific scenarios in the paper. Unfortunately, the best forum for determining its validity, publication in a scientific journal, has been effectively foreclosed by the specter of litigation, which I shall discuss below.

There are some reasons to be confident about the substantial value of the manuscript. In the time-honored tradition, its authors circulated the manuscript to others who could evaluate it from many intellectual vantage points. These included biomedical scientists, clinical and biomedical researchers, sociologists, and students of science policy, some of them members of the National Academy of Sciences. Many of these readers offered criticisms or corrections which were incorporated into the current version of the paper. I believe that Feder and Stewart conscientiously sought qualified external opinion and revised their manuscript in response to criticism. They estimate approximately 150 reviews. Over the years I, myself, have read several versions of the paper and noted many revisions and corrections. The manuscript has been sent to professional journals and has been seriously considered, and, once again, reviewed by basic and clinical researchers, by social scientists, policy experts, and some of the scientists who served on the investigative panels. It was sent to some of the coauthors whose actions were impugned.

It is not often that the subjects of research become part of the review process. This probably happened because editors were sensitive to the possibility that some of the Feder-Stewart contentions could be disputed by the coauthors who felt that their actions had been misrepresented; and one can appreciate the concern that their credibility as scientists would be damaged as a result. One must applaud the caution with which the editors were proceeding and note that Feder and Stewart took pains to respond to the coauthors' objections although they did not agree with many of them.

Reviewers typically submit their opinions to the editors, who then send them anonymously to the authors. In the case of the Feder-Stewart manuscript some of these comments were very detailed. Many were critical, many praised the manuscript. Feder and Stewart examined these criticisms and answered them, sometimes by correcting errors, sometimes by rewording or changing the emphasis. The manuscript was tentatively accepted by an extremely prestigious journal. Dates for galleys and publication were discussed. Its editor had solicited and received detailed criticism. He also received a letter from a lawyer representing a scientist whose work was criticized in the Feder-Stewart paper. This letter said that the article was "clearly defamatory" to his client and that it would cause "irreparable and serious damage" to the professional reputation of his client, and if published would be considered "published by you * * * with actual malice * * *." After several postponements of the publication date of that journal, Feder and Stewart withdrew the manuscript.

It was submitted to another significant basic science journal. This editor sent the manuscript out to many reviewers. He also received letters from lawyers representing two of the many coauthors

whose conduct was the subject of the Feder-Stewart manuscript. One of the letters stated that "publication of the paper as it presently stands will unlawfully libel and defame" his client.

In my opinion, any reasonable person who read the exchange of correspondence between the parties would come to the conclusion that fear of libel action was an important factor in the journals' decision processes about the publication of the Feder-Stewart manuscript. NIH received a request that it require that copies of the lawyers' letter and memoranda be sent to every journal where the Stewart and Feder manuscript had been submitted. Although NIH did not require they do so, Feder and Stewart had sent copies to journals to which the manuscript was formally submitted. It is likely that the high costs of potential litigation affected the decision processes of those journals as well.

Serious as they are, I believe differences of opinion about the manuscript could have been worked out in the ordinary processes of scientific review. For instance, it is possible that on the basis of broad-based scientific review, editors might have made the decision not to publish the Feder-Stewart manuscript. Or, if it had been selected for publication, I believe that a reasonable scientific journal editor would have given the Darsee coauthors ample space to reply, probably in the same issue. There is precedent for this in the case of some scientific controversies.

I deplore the intrusion of threatened legal action into the evaluation and publication processes of scientific research. There is an ominous hint of the danger in another, unrelated case of alleged fabrication where a senior author decided not to retract some dubious papers because of the possibility of a lawsuit.

Although the review processes of science are not perfect, their maintenance and improvement are extremely important. Regulation is not, I believe, a viable alternative. I would like to speak against governmental regulation of scientific research and suggest some modest tightening of supervision on the part of the granting agencies.

It is probably impossible to regulate the creative process, but that does not mean that, with a congressional mandate, people wouldn't try. It would be stultifying if there were any attempts to monitor the daily actions and professional and social interactions that surround and stimulate the genesis of new ideas and the advancement of knowledge. Furthermore, it would be extremely wasteful of a precious and increasingly scarce resource, scientific talent.

Routine internal audits are a normal means of verifying financial responsibility in most corporations. Similarly, good manufacturing practices are audited internally by pharmaceutical and medical device manufacturers to ensure the quality of their products and the safety of the public. Admittedly these internal quality review mechanisms have a mighty stimulus in the regulations of the Securities and Exchange Commission and the Food and Drug Administration. But Federal regulation of research would be counterproductive.

There are currently several kinds of informal and formal quality audit mechanisms in science. Routine laboratory reviews of data and seminars are examples. Some department chairmen examine

all manuscripts which are submitted for publication. Institutional review boards are increasingly concerned with the quality of research. Are they perfect? No. Can they be improved? Yes. A valuable adjunct to these mechanisms are the site visits conducted by granting agencies. As an initial response, stricter site visits could serve as occasions for a more comprehensive and systematic evaluation of standard laboratory practices and of the local supervision that ensures or should ensure that protocols are followed. I recommend that current means of quality assurance be strengthened and submit that many reforms are already underway.

In summary, the examination of the research and publication practices of scientists is a legitimate and important scholarly activity. The results of such studies are the appropriate concern of Federal funding agencies and the Congress. These studies are difficult to carry out because they concern sensitive and critically important issues touching on the integrity of individual scientists. In the interests of professional improvement scholars ought to be able to tolerate a good deal of mutual criticism. Research directed at correction of research practices should not be discouraged by the intrusion of defensive threats of legal action.

Second, several of the research behaviors which Feder and Stewart describe deviate substantially from generally acceptable laboratory research procedures. Some would be considered as research malpractice or seriously negligent conduct by most working scientists. Other behaviors are careless and may be misleading to other scientists. But, although they may be used opportunistically to advance careers, they sometimes happen accidentally. Suitably strengthened, internal reviews could correct those behaviors.

While a few scientists and their detractors think there is a crisis in research, a number of responsible scientists have already taken steps to reassert and revitalize the traditional values of scientific research. The Feder and Stewart study itself was conducted by scientists and relies on investigations conducted by other concerned scientists. One major area of concern is that some senior authors have been lax in their interpretation about the amount of guidance and supervision required by junior scientists, and derelict in meeting their responsibilities as joint authors on manuscripts submitted for publication in scholarly journals.

And finally, if libelous statements are made and cynically forwarded under the name of science, they should not be immune from legal recourse. However, there is an effective forum for critical evaluation in the profession—scientific journals. The maintenance of those processes of evaluation is crucial to the advancement of knowledge. Although such professional criticism is uncomfortable, even painful, for all the parties concerned—those who are the subjects of the research and, frankly, those who conduct the research—scholars cannot afford to be too thin-skinned if they are to maintain standards and preserve their professional autonomy. Whenever possible, scientists need to reinforce their own review systems and not undermine them by premature resource to legal action.

Thank you very much.

[The prepared statement of Dr. Woolf follows:]

TESTIMONY PREPARED FOR THE TASK FORCE ON SCIENCE POLICY OF THE COMMITTEE ON SCIENCE AND TECHNOLOGY, U. S. HOUSE OF REPRESENTATIVES MAY 14, 1986

Revised May 27, 1986

I am Patricia Woolf, at present a Lecturer in the Sociology Department of Princeton University. I have written and published on communication practices among scientists. For the last several years I have been studying scientific fraud and the role of journal editors in quality control in scientific research and publication. I was invited to participate in this hearing to discuss whether a manuscript describing a research project conducted by two NIH scientists, Ned Feder and Walter Stewart, contains evidence of serious problems in the way scientific research is currently conducted and communicated in America.

A brief answer is that their account is an original contribution to an important discussion. If it is accurate and representative, it clearly suggests that some reforms of research and publication practices are needed. Although serious questions have been raised about both the style and the substance of the manuscript, I believe that the processes of peer review, if allowed to function normally, were and are adequate to assess the potential value of the manuscript. Furthermore, science and the public would have been better served if these processes had been allowed to evaluate the research without the distorting effects of potential libel action.

Progress in science depends on maintaining two attitudes, scepticism and collegiality, which are not easily reconciled one with the other. No other profession makes such high demands on its practitioners in terms of the need for constant and critical scepticism about ideas, and trustful cooperation of colleagues who produce those ideas. Without collegiality and trust the complex cooperative enterprise that characterizes modern science would be slowed to a standstill. There is little doubt that collegiality can make objective criticism extremely difficult and undermine the ability to maintain healthy scepticism. But scepticism is crucial to critical evaluation of the ideas, data and conclusions of scientific research. In ordinary life scepticism undermines friendships; trust engenders friendships. Still, science demands that we dissociate our judgements about people from our acceptance of their ideas.

Collegiality depends on the assumption of integrity. A serious question for science is whether this assumption interferes with our ability to be appropriately critical of our colleagues research, and to be sufficiently diligent in detecting problems. A more serious question is how we can insure that the processes of scientific research and appraisal are sufficiently rigorous to minimize and correct errors and to eliminate fraud to the greatest possible extent.

In the past, incidents of fraud or falsification in science were rare. Among the most notable were Piltdown Man and Cyril Burt's studies of nonexistent twins. The painted mouse incident shocked the biomedical community in 1975. Then in the summer of 1980 four cases of fraud in science were disclosed and received national publicity. Since then the scientific community has set itself the task of examining and improving its quality assurance programs. It has been prompted to do so by Congressional hearings and by formal inquiries by the Federal agencies that fund scientific research. But more than those reasons, it was prompted to do so by a profound sense of outrage and embarrassment on the part of scientists. Fraud in science had the unfortunate effect of undermining trust among some scientists, but I believe it has had a compensating virtue of stimulating self-criticism. This has taken place locally at institutions where fraud has occurred and across the nation.

Responsible Authorship

It has become clear that responsible authorship of scientific papers is central to quality control in research and publication. The criteria for authorship have been appraised and definitions suggested by spokesmen representing the two major institutions of science: research universities and scientific journals. President Donald Kennedy's message, "On Academic Authorship," to the Stanford University faculty received wide attention in the scientific press and was distributed to members of the Association of American Universities. Scientific editors, among them Edward Huth of Annals of Internal Medicine and Arnold Relman of the New England Journal of Medicine have taken the lead in America in specifying contributions that scientists should have made in order to assume authorship, and the credit and the responsibilities that authorship entails. The American Chemical Society has recently published a revised version of its code of ethics which includes criteria for authorship. The Council of Biology Editors and the American

Psychological Association have written authorship policies. Unanimous agreement among the International Committee of Medical Journal Editors has produced a revision to their Uniform Requirements for Manuscripts Submitted to Biomedical Journals, a style ~~and~~ substance guide to which approximately 300 journals currently adhere.

Increased Vigilance

There is evidence, but it is anecdotal and not systematic, that there is now increased vigilance in collaborative efforts, in preservation of data and in reviewing and editing manuscripts. One of the coauthors in a well known case of fraud said that people were "tending to be more careful," especially in multinstitutional collaborative studies. New precautions are being taken with respect to data storage, data analysis and systematic record keeping. In addition, researchers have been more scrupulous in checking the credentials of potential research collaborators. Although it is not a congenial attitude, I expect that in view of recent cases of fabrication of data, co-workers and reviewers will consider the possibility of fraud along with simple error as an explanation for discrepancies in scientific data.

It would be important to document such changes in attitude and practice among research scientists. But systematic empirical studies of scientists' research and communications behavior are extremely rare. In the conduct of such studies it would be extremely important to avoid the imputation that because research practices in a given laboratory have been tightened up, the prior behavior was corrupt.

Most of the self-corrective measures taken by scientists have taken the form of examinations of conscience and re-examinations of standard practices. These appear to be based on the belief that individual integrity and collegial vigilance are the most efficient means of assuring quality and honesty in science. I think that is the case. There have been several symposia sponsored by the AAAS, such as the recent one on Pressure to Publish and Fraud (May 1985) and a continuing project on Openness and Secrecy in Science and Technology. There have been sessions of the Society for Scholarly Publishing and the Council of Biology Editors devoted to the prevention of fraud in science and ways to mitigate its effects. Scholarly research on these topics has been funded by the National Science Foundation (NSF).

The manuscript by Stewart and Feder, "Professional Practices Among Biomedical Scientists: A Study of a Sample Generated by an Unusual Event,"

has been called "the most famous unpublished paper in America." It is one of many responses to an egregious instance of fraud in science. Like other studies it has gone beyond criticism of flagrant plagiarism and falsification of data to include other misconduct in scientific research and publication such as duplicate publication, breaking the report of one experiment into several shorter manuscripts, failure to maintain careful records and raw data, and using statistical or graphical techniques to make poor data look better. These abuses appear to be for the purpose of publishing more journal articles, rather than for the advancement of scientific knowledge. They suggest that in some cases publication has become a device for professional advancement rather than a means of communicating new knowledge.

The Stewart-Feder paper is unlike other studies in one sense. It is unique in that it uses an analysis of published journal articles and reports of investigating committees to examine some of the research and publishing practices of the coauthors of those papers. It is important to note at the outset as Feder and Stewart do that the group of authors is not a random sample and may not be regarded as representative of any larger group of scientists. Furthermore, the Stewart-Feder paper does not discuss all of the publications of the scientists involved; it discusses only that research which was conducted and published in connection with a scientist who was later found to have forged scientific data.

Does the Stewart-Feder paper indicate that there is something seriously wrong in American research? First let me say that I believe their paper to be a conscientious attempt to answer a difficult question that most scientists and policy makers are interested in and most sociologists of science have found so difficult that they have not even tried to answer it. That question is "How prevalent is misconduct in scientific research and publication?" When asked whether disclosed cases of fraud are "the tip of the iceberg," sociologists shrug and reply that statistics on all kinds of deviant behavior are notably hard to come by. The Stewart-Feder paper is a serious critique of American science.

If the account is accurate it should be a matter of grave concern to scientists themselves and to Federal agencies that fund science. I am not a cardiologist and have not conducted any bench research for more than 20 years, and therefore cannot adequately judge the technical details of some of the specific scenarios in the paper. Unfortunately, the best forum for

determining its validity, publication in a scientific journal, has been effectively foreclosed by the spectre of litigation, which I shall discuss below.

There are some reasons to be confident about the substantial value of the manuscript. In the time-honored tradition, its authors circulated the manuscript to others who could evaluate it from many intellectual vantage points. These included biomedical scientists, clinical and basic researchers, sociologists, and students of science policy, some of them members of the National Academy of Sciences. Many of these readers offered criticisms or corrections some of which were incorporated into the current version of the paper. I believe that Feder and Stewart conscientiously sought qualified external opinion and revised their manuscript in response to criticism. They estimate approximately 150 reviews. Over the years I have read several versions of the paper, and noted many revisions and corrections. The manuscript has been sent to many professional journals and has been seriously considered, and, once again, reviewed by basic and clinical researchers, by social scientists, policy experts, and by some of the scientists who served on the investigative panels. It was sent to some of the coauthors whose actions were impugned.

It is not often that the subjects of research become part of the review process. This probably happened because editors were sensitive to the possibility that some of the Stewart-Feder contentions could be disputed by the coauthors who felt their actions had been misrepresented; and one can appreciate the concern that their credibility as scientists would be damaged as a result. One must applaud the caution with which the editors were proceeding and note that Feder and Stewart tock pains to respond to the coauthors objections although they did not agree with many of them.

Reviewers typically submit their opinions to the editors, who send them anonymously to the authors. In the case of the Stewart-Feder manuscript some of these comments were very detailed. Many were critical, many praised the manuscript. Feder and Stewart examined these criticisms and answered them, sometimes by correcting errors, sometimes by rewording or changing the emphasis. The manuscript was tentatively accepted by an extremely prestigious journal. Dates for galleys and publication were discussed. Its editor had solicited and received detailed criticism. He also received a letter from a lawyer representing an author whose work was criticized in the Stewart-Feder

paper. This letter said that the article was "clearly defamatory" to his client and that it would cause "irreparable and serious damage" to the professional reputation of his client, and if published would be considered "published by you and the authors with actual malice . . . " After several postponements of the publication date, Feder and Stewart withdrew the manuscript.

It was submitted to another significant basic science journal. This editor sent the manuscript out to many reviewers. He also received letters from lawyers representing two of the many coauthors whose conduct was the subject of the Stewart-Feder manuscript. One of the letters stated that "publication of the paper as it presently stands will unlawfully libel and defame" his client. In my opinion, any reasonable person who read the exchange of correspondence between the parties would come to the conclusion that fear of libel action was an important factor in the journals' decision processes about publication of the Stewart-Feder manuscript. NIH received a request that it require that copies of the lawyers' letters and memoranda be sent to every journal where the Stewart and Feder manuscript had been submitted. Although NIH did not require they do so, Feder and Stewart had sent copies to journals to which the manuscript was formally submitted.

Serious as they are, I believe differences of opinion about the manuscript could have been worked out in the ordinary processes of scientific review. For instance, it is possible that on the basis of broad based scientific review, editors might have made the decision not to publish the Stewart-Feder manuscript. Or, if it had been selected for publication, I believe that a reasonable scientific journal editor would have given the Darsee coauthors ample space to reply, probably in the same issue. There is precedent for this in the case of some scientific controversies. I deplore the intrusion of threatened legal action into the evaluation and publication processes of scientific research. (There is an ominous hint of the danger in another unrelated case of alleged fabrication where a senior author decided not to retract some dubious papers because of the possibility of a lawsuit.)

Although the review processes of science are not perfect, their maintenance is extremely important. Regulation is not a viable alternative. I would like to speak against governmental regulation of scientific research and suggest some modest tightening of supervision on the part of granting agencies.

WOOLF 5/14/86

It is probably impossible to regulate the creative process, but that does not mean that, with a Congressional mandate, people wouldn't try. It would be stultifying if there were any attempts to monitor the daily actions and professional and social interactions that surround and stimulate the genesis of new ideas and the advancement of knowledge. Furthermore it would be extremely wasteful of a precious and increasingly scarce resource, scientific talent.

Routine internal audits are a normal means of verifying financial responsibility in most corporations. Similarly, good manufacturing practices are audited internally by pharmaceutical and medical device manufacturers to insure the quality of their products and the safety of the public. Admittedly these internal quality review mechanisms have a mighty stimulus in the regulations of the Securities and Exchange Commission and the Food and Drug Administration. But federal regulation of research would be counterproductive.

There are currently several kinds of informal and formal quality audit mechanisms in science. Routine laboratory reviews of data and seminars are examples. Some department chairmen examine all manuscripts which are submitted for publication. Institutional review boards are increasingly concerned with the quality of research. Are they perfect? No. Can they be improved? Yes. A valuable adjunct to these mechanisms are the site visits conducted by granting agencies such as the NSF and NIH. As an initial response, stricter site visits could serve as occasions for a more comprehensive and systematic evaluation of standard laboratory practices and of the local supervision that ensures that protocols are followed. I recommend that current means of quality assurance be strengthened and submit that many reforms are already under way.

Summary Thoughts

1) Examination of the research and publication practices of scientists is a legitimate and important scholarly activity. The results of such studies are the appropriate concern of federal funding agencies and the Congress. These studies are difficult to carry out because they concern sensitive and critically important issues touching on the integrity of individual scientists. In the interests of professional improvement scholars ought to be able to tolerate a good deal of mutual criticism. Research directed at correction of research practices should not be discouraged by the intrusion of defensive threats of legal action.

2) Several of the research behaviors which Feder & Stewart describe deviate substantially from generally acceptable laboratory research procedures. Some would be considered as research malpractice or as seriously negligent conduct by most working scientists. Other behaviors are careless and may be misleading to other scientists. But, although they may be used opportunistically to advance careers, they sometimes happen accidentally . Suitably strengthened, internal reviews could correct those behaviors.

3) While a few scientists and their detractors think there is a crisis in research, a number of responsible scientists have already taken steps to reassert and revitalize the traditional values of scientific research. The Feder & Stewart study itself relies on investigations conducted by concerned scientists. One major area of concern is that some senior authors have been lax in their interpretaion of the amount of guidance and supervision required by junior scientists, and derelict in meeting their responsibilities as joint authors of manuscripts submitted for publication in scholarly journals.

4) If libelous statements are made and cynically forwarded under the name of science, they should not be immune from legal recourse. However, there is an effective forum for critical evaluation in the profession - scientific journals. The maintenance of those processes of evaluation is crucial to the advancement of knowledge. Although such professional criticism is uncomfortable, even painful for all the parties concerned, scholars cannot afford to be too thin-skinned if they are to maintain scholarly standards and professional autonomy. Whenever possible, scientists need to reinforce their own review systems and not undermine them by premature recourse to legal action.

WOOLF 5/14/86

DISCUSSION

Mr. FUQUA. Thank you, Dr. Woolf.

You mentioned in the course of your testimony that some of the scientific societies and professional organizations have taken steps to strengthen their code of ethics as it relates to authorship, and so forth. Do you see that emerging as a greater awareness in recent times to try to correct what might be perceived as some of the problems that were outlined in the Feder and Stewart report?

Dr. WOOLF. Yes, I do. A lot of the focus has been on responsible authorship and has reaffirmed that if authors are to gain the benefit of being an author of a paper they also have to take responsibility for what is in that paper.

Mr. FUQUA. You likewise mentioned that Federal regulation, or agency regulation was not necessarily the answer and I agree. But at some point there must be some self-governing to make sure that we do not even give the hint of condoning or permitting this type of activity to continue. And I gather you are saying that self-discipline through various organizations and the university research groups would be the most effective method of trying to cope with this malpractice.

Dr. WOOLF. I think they are the most effective. I think we can do better. I think inquiries of this kind quicken the mind and make people attentive to the real seriousness of these practices. I deplore the facile writing off misconduct as little white scientific lies. We have to set a very, very high standard.

Mr. FUQUA. Do you think this is widespread in the scientist, little white scientific lies or big lies, or whatever mechanism?

Dr. WOOLF. I think that there has been some opportunistic exaggeration of the results of research in order to make more of a publication than the research may warrant.

Mr. FUQUA. The publish or perish syndrome.

Dr. WOOLF. Well, I think that is important. But I think it is also important to recognize that pressure to publish is not a ubiquitous feature of American science. I believe that in most places there is a reasonable response to a reasonable amount of pressure. That is, if the granting agency and the American people support research, there ought to be some response to that. There ought to be felt an appropriate amount of pressure to produce a product that was paid for by the American people.

What troubles me is that in some places there is an exaggerated response to that, and a level of publication which by simple time and motion studies would lead you to doubt that it could have been adequately and carefully done.

Mr. FUQUA. Do you see, though, that the system where it is publish or perish or your standing in the department is based on how much money you bring in and you bring in money by getting grants, and that affects your tenure, your standing as far as promotion within the faculty, or it can have an impact on that, and that the system is the driving force and the system tends to pressure people to publish and many are under immense pressure to succeed or to advance, that the system pushes people to involve themselves in this activity?

Dr. WOOLF. I think the system works well in most places, in most institutions, in most science departments across the country, and the publication level of most scientists in most departments would indicate that that is the case. Typically, the average NIH grant, for instance, taking the number of grants divided by the number of published papers you get about 1.3 papers per grant. Even if these are unevenly distributed, that is not an unreasonable expectation for the level of publishing we ought to expect for public investment.

However, there are pockets of extravagant, prolific publishing in which I think there is excessive pressure to publish and that it is responsible for causing or allowing some of this kind of misconduct.

Mr. FUQUA. We discussed this in some of our earlier hearings in the task force, about how do you measure academic excellence and is it based on the amount of money that you bring into the department in the way of research grants or is it how you communicate and teach or your ability to work with your students and advance them? That is always a question that, when you talk to the administrators of the department, they are the ones that are probably looking for the money, and it puts somewhat of an inherent conflict and a pressure to publish; and therefore in publishing you may attract more research money. And I think we all recognize that society has its imperfections and that is not just limited to ordinary life but also to the universities.

I guess I am not asking that in the form of a question, but it is something that we have discussed here. And I would be interested in any comments that you might have on that.

Dr. WOOLF. I think that it is important that the tail not be allowed to wag the dog and that the search for grant money to conduct research not be allowed to distort the research that is produced with those public moneys. The danger of competition I think would always be with us. It is important to be wary, and I think as a result of these recent actions and as a result of the kind of inquiries that are being made people are more wary.

Mr. FUQUA. Do you have any comment on some of the action that some of the Federal agencies have taken such as NIH? You were here when I asked a similar type question to Drs. Feder and Stewart about appropriateness of responses from the Federal to try to discourage and, in effect, eliminate this type of practice.

Dr. WOOLF. I think I agree with Drs. Feder and Stewart that the NIH investigative panel did a very good job.

Mr. FUQUA. I am speaking also of some of the other agencies such as NSF and the other granting agencies, too.

Dr. WOOLF. I made specific inquiry about that situation at the National Science Foundation and was told that they perceive that there is not yet a serious problem and they think it is an unwise use of funds to gin up the solution to a problem that has not yet forwarded itself. I think that my initial points about collegiality and skepticism make it very difficult for scientists and for the granting agencies, many of which are governed by scientists, to suspect deep flaws in their processes. We have been shocked into that recognition and I think are much more attentive to it than we used to be.

I think that due process in these situations is much better now and likely to be much better if there are any examples of this in the future. Precedents have been set, but at the beginning, as Drs. Feder and Stewart indicated, there was a good deal of fumbling around and a good deal of denial. A great deal of unsurety as to who should be notified. A good deal of secrecy. It wasn't at all clear how much the public ought to know or how much people thought was appropriate for the public to know in these situations as they first emerged.

Mr. FUQUA. Do you think the peer review process as we see it today, or as it operates today, should have an additional responsibility of trying to make sure that, not that they can maybe read all the publications that have been published by a proposal or chief investigator, but that it should be a part of their responsibility to make sure that these people are for real and not having bogus publications?

Dr. WOOLF. I think that that should be an obvious and important part of peer review. And I would not say that it always functions as well as it might. There is a recent book on peer review by Stephen Lock of the *British Medical Journal.* A number of articles on improving peer review have been published. Still, there is a great deal of pressure on scientists when they review manuscripts, and their responsibility in that case is a corporate responsibility, a professional responsibility, one that does not improve necessarily their own status in a institution. Perhaps more recognition should be given to the amount of time that good scientists do spend reviewing manuscripts, and more professional advancement for that part of their professional work.

Mr. FUQUA. Thank you very much.

Dr. WOOLF. Thank you very much, Mr. Chairman.

[Whereupon, at 11:50 a.m., the task force was adjourned, to reconvene on Thursday, May 15, 1986.]

APPENDIX 1

STATEMENT FOR THE RECORD BY
BANCROFT LITTLEFIELD, JR., ESQUIRE
FOLEY, HOAG & ELIOT
BOSTON, MASSACHUSETTS

AND RESPONSE BY WALTER W. STEWART
AND NED FEDER
NATIONAL INSTITUTES OF HEALTH, BETHESDA, MARYLAND

STATEMENT FOR THE RECORD

RE: Response To The May 14, 1986 Statement of Walter W. Stewart and Ned Feder Before The Task Force on Science Policy of the Committee on Science and Technology, U.S. House of Representatives

FROM: Bancroft Littlefield, Jr., Esquire
Foley, Hoag & Eliot
One Post Office Square
Boston, Massachusetts 02109

DATE: September 18, 1986

INTRODUCTION

In a May 14, 1986 written statement to the Task Force on Science Policy, Walter W. Stewart and Ned Feder reported the results of their unpublished paper, "Professional Practices Among Biomedical Scientists: A Study Of A Sample Generated By An Unusual Event." That paper, a copy of which was submitted with their statement to the Task Force, purports to examine the professional practices of 47 biomedical scientists who had coauthored publications with John Darsee, a researcher found by the National Institutes of Health to have engaged in widespread scientific fraud at two universities during a period of over twelve years. Although Mr. Stewart and Dr. Feder emphasize that none of the 47 coauthors knew of, authorized, or otherwise condoned Dr. Darsee's fraudulent acts, Mr. Stewart and Dr. Feder contend that their study raises "the possibility that misconduct may be far more common among biomedical scientists than is realized." (Statement at p. 5).

Undoubtedly, the potential costs of serious misconduct in research are enormous, not only for the persons and academic institutions involved, but for science as an institution. A careful and accurate analysis of research practices which offered some insight into the frequency of scientific misconduct and the manner in which it could be prevented or detected would be an important contribution to the public understanding of this issue and the formulation of public policy. It is therefore regrettable that Mr. Stewart and Dr. Feder have missed this opportunity to make such a meaningful contribution both in their statement to the Task Force and in their paper. Instead, Mr. Stewart and Dr. Feder have provided a paper which contains inaccurate and unsupportable allegations with respect to the Darsee matter and publications coauthored by Darsee.

In their paper, Mr. Stewart and Dr. Feder claim to have identified a number of so-called questionable practices in several of the Harvard research papers and abstracts on which Dr. Darsee was a coauthor. It is the purpose of this memorandum to review the three principal allegations in the paper regarding those Harvard publications. The points raised in this Statement have previously been communicated to Mr. Stewart and Dr. Feder during the review process of their paper. When examined closely, each of the three allegations in the paper proves to be rooted in an inaccurate depiction or interpretation of the operative facts or to arise from legitimate differences of scientific opinion. There was no "unacknowledged republication" of data from one Harvard paper in a second Harvard paper authored by the same

scientists because the second paper expressly referred twice in its opening paragraph to the previous paper and whatever repetition occurred was necessary to provide a complete analysis of the research reported on in the second paper. While the Harvard coauthors acknowledged that more complete detail regarding the source of control data in several papers could have been provided, the practice in the laboratory regarding control data was entirely appropriate, and the absence of this detail in these papers is no more than a difference in scientific opinion because it affected neither the validity of the results nor their reproducibility. The so-called minor "errors" Mr. Stewart and Dr. Feder purport to have found are, for the most part, not errors at all but arise from either scientific misunderstandings of Mr. Stewart and Dr. Feder, efforts to extrapolate precise values from figures or graphs in the papers or amount, essentially, to typographical errors. These allegations, which are the ones directed to the research papers from Harvard, do not sustain the conclusions reported by Mr. Stewart and Dr. Feder.

Indeed, it was the inaccuracy of allegations made by Mr. Stewart and Dr. Feder that prompted John Maddox, Editor of Nature, to write to them in January, 1985 that Nature would not publish their paper unless certain specific corrections were made. These requested corrections included, inter alia, removing allegations that sought to impute by speculative inference "motives or states of mind to the authors...." According to Mr. Maddox, the facts would not support the charge that Darsee's coauthors had engaged in intentional and knowing acts of

dishonesty, even if that charge was toned down to speak of an "apparent lack of candor": "the phrase 'apparent lack of candor' is indistinguishable from 'dishonesty'. I do not believe it can be sustained, nor do I think it to be necessary." (emphasis added)

In a postscript, published in the March 6-12, 1986 issue of Nature, Mr. Maddox responded to public statements by Mr. Stewart and Dr. Feder that comments to two journals by attorneys representing Harvard coauthors had blocked publication of the paper. He wrote:

> "Strictly speaking, it is not correct that some of the unnamed but identifiable co-authors threatened Nature with a libel suit, but they, and their attorney, argued that the manuscript in its original form and in many later versions would have been unwarrantably damaging. This is true, and has been accepted to be so." (emphasis added)

Over the past two years, Mr. Stewart and Dr. Feder have submitted their paper for publication to at least 15 separate scientific journals. In each case, however, the journal found their paper unfit to publish. Some journal editors believed their journal to be inappropriate for an article of this nature; others questioned the validity of the conclusions the paper reached or disagreed with important parts of the paper; still others found the paper as a whole to be flawed. For example, Dr. Singer wrote:

> "The subject matter does not seem to me appropriate for the proceedings of the National Academy of Sciences. It does not, in my judgment, meet the general standard of original theoretical or experimental research. The paper, by its own statements, is an evaluation of a set of data against certain standards which the two of you have selected. Thus I would characterize the

paper as a compilation and analysis of data and a critique, not theoretical or experimental research. It is beside the point whether or not I share the standards you bring to the critique or agree with your conclusions."*

Steven Harnard, Editor of The Behavioural and Brain Sciences, wrote:

"...I believe that there are important sampling problems that make it difficult to draw conclusions from these data. ..."

"The basis for your sample was co-authoring with Darsee. This leaves it completely indeterminate whether or not this population is representative of the larger biomedical community and its practices. Notice that one can draw neither conclusion: neither that the practices are typical, and represent a general problem, nor that they are specific to Darsee collaborators and represent a specific case only. Related to this is the problem of establishing a base line for these various practices, to determine what constitutes standard and what constitutes deviation (and how much) and when. There is also the issue of the relative severity of these practices. Some are potentially more important than others; and all (as I suggested) ought to be weighted by the importance of the research itself (which in many cases is not that great). These two are sampling and normative questions that would require a much more systematic procedure to answer. Finally, the damage that the report would do to the investigators in question would perhaps be justified if these were normative results, based on representative, interpretable samples. But, as I have argued, they are not; so the potential injustice in this case becomes quite a sizable factor."*

---

* Correspondence relating to the paper by Mr. Stewart and Dr. Feder was made available by the NIH under the Freedom of Information Act to The New York Times in March, 1986, and subsequently in May, 1986 to the Harvard coauthors.

Even more critical of the paper by Mr. Stewart and Dr. Feder was Thomas P. Stossel, M.D., Editor, Journal of Clinical Investigation:

> "I found the paper to be overstated, woolly, and excessively judgmental and moralistic."*

RESPONSE TO THE SUBSTANTIVE ALLEGATIONS

Taken at face value, some of Mr. Stewart and Dr. Feder's allegations and "evidence" would appear to raise real and substantial issues regarding research practices which seemingly depart from generally accepted standards. To the casual reader, or indeed any reader who does not have the lengthy time required to make an independent comparison of the allegations with the specific evidence in the Harvard papers, the claims of Mr. Stewart and Dr. Feder may appear significant. But as have concluded reviewers and editors who have gone through the charges point by point and measured them against the concrete evidence, Mr. Stewart and Dr. Feder's assertions involving the Harvard papers simply are not borne out by that evidence.

Essentially, three areas of criticism of the Harvard publications are raised in Mr. Stewart and Dr. Feder's paper. One criticism, that involving the alleged failure to describe precisely that some control-animal data were being reused in experiments, is one that if it were accurately defined would be, in the Harvard coauthors' view, a minor but legitimate criticism

---

* Correspondence relating to the paper by Mr. Stewart and Dr. Feder was made available by the NIH under the Freedom of Information Act to The New York Times in March, 1986, and subsequently in May, 1986 to the Harvard coauthors.

stemming from a difference in scientific opinion. The second, involving the Harvard coauthors' alleged republication of data without acknowledgement, is not, in the Harvard coauthors' view, a legitimate criticism. The third area of criticism, that of readily detectable, though minor, errors in the tables, diagrams, and text of the Harvard papers has been described by Mr. Stewart and Dr. Feder as a less serious assertion. Moreover, a review of the fifteen so-called errors that Mr. Stewart and Dr. Feder have specifically identified reveals that even that less serious criticism cannot be sustained: Eight of the purported errors are not errors at all. Three supposed errors are traceable to Mr. Stewart and Dr. Feder's attempt to extract from a diagram or graph a precise numerical value which in turn is compared to the precise numerical value listed in the text. Two other alleged errors are obviously minor typographical errors and two are minor errors which do not mislead the reader or affect the conclusions of the paper. In no event are these alleged errors violations of accepted scientific research practices and by no stretch of the imagination do they amount to "scientific misconduct" as claimed by Mr. Stewart and Dr. Feder.

The following is a point-by-point discussion of each of these criticisms.

Criticism No. 1 -- "Misleading Statements Regarding the Descriptions of Experimental Methods in Five Harvard Papers."

The challenged statements relate to the alleged use of "historical" rather than contemporaneous controls and the failure adequately to describe this practice in the published papers from Harvard.

The experiments involved subjecting a set of dogs to coronary occlusion and then intervening with various treatments and studying the resulting effect on myocardial infarct size. After occlusion, the dogs were randomly assigned to control and treatment groups. The treated dogs were then subjected to some form of intervention, and after the relevant time period, both the control and treated dogs were sacrificed to enable data collection and comparisons between the treated and control groups. For this first experimental intervention, the controls were fully contemporaneous.

To study the effect of a different treatment, new dogs were then subjected to coronary occlusion. At this point, the investigators had several choices with respect to control data. At one extreme, they could have used entirely new control dogs for each new intervention. This choice, of course, would have resulted in the sacrifice of the largest number of animals, for most data could not be collected without sacrificing the animals. At the other extreme, the Harvard coauthors could have used the control-data information they had collected when the control dogs from the first intervention had been sacrificed and their relevant parameters measured. This choice might have introduced

error had there been variations over time between the previous controls and the baseline from which the treated dogs in later experiments were measured.

To avoid this potential variability while minimizing the unnecessary sacrifice of dogs, the Harvard coauthors chose a third and middle course. For the second and subsequent interventions, they studied some control dogs de novo, i.e., contemporaneously with the treated dogs, but allowed the remainder to be selected at random from among the control data collected previously. They allowed this practice only when the research fellow who performed the first and the subsequent intervention was the same and when the experimental conditions were similar. Moreover, to assure that contemporaneous controls yielded reliable data, the investigator was to check them against the data yielded from the previous controls. If the current and previous controls were found to yield consistent information, they stopped short of sacrificing an entire new group of control dogs for each intervention.

It must be emphasized that the use of previous control data was not authorized unless a proper, random selection of that data was made. The coauthors' view was that, had this practice been employed as specified, it would not have differed materially from the use of entirely new control dogs at each intervention. This practice did not involve the use of "historical" controls as that term is commonly understood, for the experiments were all done within a relatively short period of time of each other. Mr. Stewart and Dr. Feder are incorrect when they allege that the

practice in the Harvard laboratory "produces a less reliable standard of comparison than controls chosen by a randomization procedure at the time of the experiment," since no controls were to be used that had not been chosen by a randomization procedure both originally when a dog was assigned to a treated or control group and in a subsequent intervention when control-dog data were randomly reused.

The Harvard coauthors have acknowledged from the beginning of the NHLBI investigation of the Darsee matter that it would have been preferable to have specified in each publication the exact details of this practice concerning control dogs. However, the coauthors included that information which they believed to be critical to another investigator's ability to assess or repeat the experiment. The precise details of how the control dogs were selected, like other variables in these experiments such as the season, the breed, age, and sex of the animals, were not believed to fall into this category. Given the emphasis placed currently on condensing submitted articles as much as possible, the coauthors judged the precise details of the control-selection procedure to be one of those details they could afford to omit. While it would have been preferable to have been more specific in these descriptions, the failure to do so hardly rises to the level of scientific misconduct, dishonesty, or unethical practice.

With hindsight, it has now been established that Dr. Darsee <u>misused</u> the authorized practice by selecting previous control data not randomly but rather with a bias toward maximizing the

variation between the control and the treated groups. But the NHLBI panel specifically found that Dr. Darsee alone was responsible for this bias, and Mr. Stewart and Dr. Feder expressly state in their paper that they agree with this finding. Nonetheless, they assert that the authorized practice was at least partially responsible for Dr. Darsee's abuse of the process. This assertion is plainly unwarranted since the Harvard coauthors reasonably relied on Dr. Darsee to adhere to the authorized practice. At the time, the coauthors had no basis whatsoever to doubt Dr. Darsee's integrity and trustworthiness as a researcher.

Finally, Mr. Stewart and Dr. Feder state that the Harvard coauthors, in letters to the NIH, "approved" of the reuse of control data. To the extent this statement refers to the randomized reuse of control data -- the policy we have described above -- it is accurate. But to the extent it is meant to, or does, suggest that the Harvard coauthors approved of Dr. Darsee's unauthorized, biased selection of previous control data, it grossly misrepresents the position taken by the Harvard coauthors in their letters to the NIH. These letters expressly criticized Dr. Darsee's failure to be "scrupulously honest" in his adherence to the Laboratory's policies regarding reuse of control data. The Harvard coauthors never approved of nor defended the way in which Darsee reused control-animal data. Indeed, they have condemned consistently that abuse of the Laboratory's practice.

To assert from these facts that the Harvard coauthors "knew or should have known" that they were making inaccurate statements in their papers, that this practice was one of the "most striking of [Mr. Stewart and Dr. Feder's] findings", and that it is an instance of scientific misconduct, is unsupportable and unfair.

Criticism No. 2 -- "Unacknowledged republication by the Harvard coauthors in a second paper of data published in a previous paper by the same authors."

Mr. Stewart and Dr. Feder claim to have "found several instances in which the same data were republished without adequate reference to that circumstance." As evidence for this statement, Mr. Stewart and Dr. Feder discuss two papers (Papers 95 and 96 in their bibliography). They then suggest that the second paper might have been "intended to conceal the connection between the two papers" and that statements in the second paper regarding the first were "misleading or inaccurate". The other instances of alleged republication involve several sets of abstracts.

Regarding the two papers, the very first paragraph of the second paper completely refutes any suggestion that the second paper might have "intended to conceal" its connection to the prior paper. The first two sentences in the second paper clearly and explicitly refer the reader to the first paper. (Reference no. 3 is the previous paper):

> "Previous studies from this laboratory have shown that brief (15 min.) coronary occlusions followed by reperfusion do not cause necrosis but do result in prolonged abnormalities in myocardial biochemistry, function and ultrastructure (1-3). At 72 hours after

> release of a 15-minute coronary occlusion, the ATP concentration in reperfused previously ischemic myocardium was significantly reduced (by 22% of normal), the percentage of systolic shortening of reperfused left ventricular segments was reduced (by 42% of normal), and ultrastructural abnormalities were present despite the absence of necrosis (1, 3). However, it was not known if these changes ever were reversed to normal after longer periods of reperfusion and, if so, what the time course of the recovery is."

As this opening paragraph makes clear, no charge of "unacknowledged republication" can be sustained.

Instead, Mr. Stewart and Dr. Feder's complaint seems to be that the data in the second paper come from the same sets of experiments underlying the first, with the first paper reporting on results after three days and the second on results obtained after seven and fourteen days. Mr. Stewart and Dr. Feder's real complaint thus appears to be that the two papers <u>could</u> have been published as one, had the authors elected to collect and analyze all the data before publishing.

The papers concern a series of experiments measuring the detailed effect of reperfusion of varied time intervals in dogs subjected to transient coronary occlusion. The work underlying these papers was designed to provide intensive and comprehensive assessments of various parameters during several periods of reperfusion following transient coronary occlusion. Some of the examination techniques yielded rapid results; others, specifically electron microscopy and histochemical analysis, required extensive examination and analysis before results could be compiled. Paper 95 reported data up to and including three days of reperfusion.

The analysis of data at three days resulted in the important finding that recovery of the heart had not occurred at that point. Paper 96 reported results for longer reperfusion periods of seven and fourteen days. Complete results for the short reperfusion periods of three days and less were obtained substantially in advance of the results for the longer reperfusion periods of seven and fourteen days.

The question of when and how best to report these results was one of scientific and medical judgment. The Harvard coauthors believed that the data obtained after three days contained surprising findings of immediate scientific interest with important clinical implications. Given recent uses of innovative therapies to open up coronary occlusions during the acute phase of a heart attack, the three-day findings suggested that such therapy might not lead to immediate recovery of heart function. Moreover, until it was learned that full recovery did not occur at three days, there was no reason to undertake seven and fourteen day studies.

The three-day findings were believed to be of significant importance in light of the rapidly developing uses of these innovative therapies in humans for coronary occlusions. But substantial additional time was required before data from the seven and fourteen day reperfusion periods could be analyzed. The Harvard coauthors therefore judged it preferable to release immediately the three-day data and to publish the seven- and fourteen-day data as they became available.

When the second paper was submitted -- some six months after the first was submitted -- it recapitulated, in the clearly identifiable form noted above, the findings at three days in order that the entire experimental picture would be complete. There was nothing improper or even unusual in the decision of the Harvard coauthors to release important new data immediately and to release subsequent data as they became available.

Moreover, the second paper presented new data on both the original twenty-one dogs for whom three-day reperfusion data had been reported and for twenty additional dogs that were studied for seven- and fourteen-day reperfusion periods. The first paper, Paper 95, reported on successful experiments on twenty-one separate dogs. The second paper reported on the results of successful experiments on twenty additional dogs that were studied for either seven or fourteen days following reperfusion, including biochemical, physiological, electron microscopic and histochemical data. The second paper also reported for the first time (1) physiological data obtained seven and fourteen days following reperfusion in seven dogs for whom data obtained 72 hours after reperfusion had been described in the first paper and (2) histochemical data on dogs sacrificed at 72 hours. It is inaccurate, in short, to allege that the two papers reported on the same data.

Mr. Stewart and Dr. Feder's speculation that the Harvard coauthors were actually motivated in the first paper by a desire to conceal subsequent republication of that same data in a paper not yet written is inaccurate.

In addition to the allegation about Papers 95 and 96, Mr. Stewart and Dr. Feder make an additional allegation concerning so-called republication of several sets of abstracts. According to Mr. Stewart and Dr. Feder, it is a questionable practice for authors to revise the title or text of an abstract upon resubmitting it for presentation at a different meeting. Mr. Stewart and Dr. Feder's view on this point defies both generally accepted standards and common sense. According to this view, once an abstract had been submitted but rejected, it could not in the future be improved, altered, tailored to the subject of a second meeting, or submitted in revised form to a national meeting after presentation at a regional meeting. That would impose on abstract writers a standard of draftmanship far too demanding of even the finest literary craftsmen.

Mr. Stewart and Dr. Feder identify three pairs of abstracts as constituting questionable republication (nos. 70 & 78, 72 & 80, and 73 & 77). All involved the submission of abstracts for presentation at meetings sponsored by the American Heart Association or the American Federation for Clinical Research. With respect to each of the three sets of abstracts, the rules of these sponsoring organizations clearly sanctioned the resubmission of those abstracts, for none was submitted for duplicate presentation at <u>national</u> meetings. Given that the sponsoring organizations authorized such resubmission of abstracts, Mr. Stewart and Dr. Feder's claim that such resubmission violated generally accepted standards is wholly without merit.

A case in point is Mr. Stewart and Dr. Feder's claim illustrated by Abstract 70, which was not accepted for presentation at the national meeting of the American Federation for Clinical Research. After revision, it was accepted as Abstract 78 for presentation at the national Scientific Session of the American Heart Association. The scientific community would not find fault with the revision of a rejected abstract designed to enhance its quality and clarity.

Mr. Stewart and Dr. Feder's claims with respect to the other pairs of abstracts are likewise mistaken. The actual facts are that one abstract in each pair (#72 and #73) was presented at the eastern regional meeting of the American Federation for Clinical Research, and the other abstract in the pair (#80 and #77) was presented at the national meeting of the American Heart Association. Hence, each of these presentations was sanctioned by the rules of the sponsoring organizations for submission of abstracts. The differences in these abstracts trace in large part to tailoring or improving the abstracts for national meetings and to the physical differences in the abstract forms supplied by the two sponsoring organizations. The form supplied for the American Federation for Clinical Research regional meeting, to which Abstract 73 was submitted, has more space for text than the form supplied for the American Heart Association meeting. Hence, Abstract 73 is longer and more detailed than Abstract 77. The same point is true for Abstracts 72 and 80. An attempt to improve an abstract by revising it or its title can hardly be interpreted as an attempt to "disguise" republication.

Criticism No. 3 -- "So-called Errors or Inconsistencies."

Mr. Stewart and Dr. Feder's paper purports to calculate the frequency of questionable practices in publications coauthored with Dr. Darsee. Yet, less than 20% of these alleged instances of questionable practices are specifically identified in the text; the rest exist as nothing more than totally unsubstantiated assertions in tables appended to the paper. Without knowing what these alleged instances of questionable practices are, or where they supposedly appear in a publication, it was simply impossible to determine whether errors in fact were committed or whether Mr. Stewart and Dr. Feder had misunderstood or mischaracterized the data in the papers.

Mr. Stewart and Dr. Feder have repeatedly declined to honor requests by the Harvard coauthors to discuss and review the specific allegations of errors and questionable practices which are not identified in the text of the paper.[1] This refusal to

---

[1] Mr. Stewart and Dr. Feder's refusal to disclose the alleged instances of questionable practices reported in the tables to their paper is contrary to the position outlined in the 1983 report of the Committee on National Statistics of the National Research Council:

> Recommendation 1: "Sharing data should be a regular practice";
>
> Recommendation 2: "Investigators should share their data by the time of publication of initial major results of analyses of the data except in compelling circumstances";
>
> Recommendation 10: "Journal editors should require authors to provide access to data during the peer review process."

The reasoning for this last recommendation is stated as follows:

(footnote continued)

disclose the substance of these unverifiable allegations of error is particularly troubling in light of the fact that, of the 15 alleged errors in the Harvard publications identified by Mr. Stewart and Dr. Feder, nearly all turned out upon careful review not to be erroneous at all. Of those errors they have specifically explained, two-thirds are not errors at all, and others turn out to be simply typographical errors or to reflect the limitations of artwork. As an analysis of these so-called errors shows, most scientists would not consider them to constitute significant departures from accepted practices.

(1) Paper 92: The following matters, identified by Mr. Stewart and Dr. Feder as errors in Paper 92, are in fact not errors, nor are they matters that in each separate instance amount to violations of generally-accepted scientific practices:

(a) Allegation that it was an error and a questionable practice for the Harvard authors, in Paper 92, to state that Figure 9 showed only a small improvement in regional function in the subepicardial segments by the end of 72 hours of reperfusion. This

---

(footnote continued from previous page)

> "Access to data during the review process, a practice already in use by some journals, provides reviewers an opportunity to replicate the analysis and discover possible errors. Reviewers can use alternate assumptions or analytic models to test the robustness of authors' conclusions."

National Academy Press, Sharing Research Data 25, 26, 30 (Feinberg, Martin, Straf, eds.) (1983).

allegation is without merit since it is based on Mr. Stewart and Dr. Feder's misunderstanding of the scientific principles involved. In this particular study, the authors interpreted recovery from zero as the appropriate baseline to gauge active function of the heart. Mr. Stewart and Dr. Feder confuse differences in interpretation with "practices which depart from generally accepted standards."

(b) Several allegations that it was a violation of generally-accepted scientific practices to publish papers with alleged discrepancies between values specified in the text and those depicted in graphic displays and figures. In each instance, the discrepancy is the result of Mr. Stewart and Dr. Feder's attempt to contrast precise values stated in the text with values portrayed in the figures. These figures reflect the numbers in the text within the technical limitations and purposes of artwork in scientific papers. In Paper 92, for instance, Mr. Stewart and Dr. Feder complain that the text states that ultrasonic crystals were placed 3mm from necrotic tissue, whereas Figure 10 appears to show a crystal touching that tissue. Figure 10 was a schematic drawing reduced many times from actual size; no reader would take that drawing as a literal depiction and be misled.

(c) Also in Paper 92, Mr. Stewart and Dr. Feder find a discrepancy between an SEM value in the text and the same value in Figure 9. That figure reflects the value within the limitations of the artwork.

(d) Allegations that the error bars in Paper 92 are implausibly similar. For this matter to be the basis of an allegation of a violation of a generally-accepted scientific practice, Mr. Stewart and Dr. Feder must claim that the authors should have recognized that these error bars were suspicious and therefore possibly inaccurate. It is untenable to claim that the similarity in the error bars was so pronounced as to be inherently suspicious. Indeed, Reference 92 was published in a peer-reviewed journal, and no questions were raised concerning those error bars. With the advantage of hindsight, we now know that much of Darsee's data were fabricated; but without that knowledge, it was not a violation of generally-accepted practice to publish the paper with the error bars in the figure.

(2) Paper 96: The following occurrences, identified (in earlier drafts and counted in the tables to the current draft) as errors in Paper 96, are in fact not errors.

(a) Allegation of a discrepancy in Paper 96 between the legend of Figure 3 which states recovery from systolic shortening did not occur until after 7 days of reperfusion and Table 1 which shows recovery at

7 days. According to Mr. Stewart and Dr. Feder, the legend to Figure 3 is "at least unclear, and we regard it as erroneous." While the phraseology could have been more artful, the difference between recovery "after 7 days" versus "at 7 days" is semantic and of the most minor significance. Indeed, given that the paper makes clear that the data displayed in Figure 3 was collected at the end of 7 days, no reader would have been misled by use of the phrase "after 7 days".

(b) Alleged discrepancy in Paper 96 between a value repeated in the summary and the text with respect to ATP concentration. It may well be that Mr. Stewart and Dr. Feder did not consider that there are at least two ways of calculating these figures and each would arrive at slightly different percentages: one involves taking the mean values for all animals and dividing ischemic by nonischemic to reach one mean percent figure, and another involves calculating the percentage on a dog-by-dog basis and then calculating an overall mean for all animals. Mr. Stewart and Dr. Feder have assumed the calculation for the summary was made by the former method.

(c) Alleged error in Paper 96 for the coauthors not to have detected the inconsistency between dp/dt values in the different publications. That allegation is based on hindsight. The inconsistency in dp/dt values is not readily discoverable. It was in

accordance with generally-accepted standards for the Harvard authors to accept Darsee's calculation of these simple values. Any other conclusion would require each coauthor of a paper to recalculate each and every numerical value.

(3) A telling example of the inaccuracy in Mr. Stewart and Dr. Feder's tables emerged from the following set of facts. In the first draft of Mr. Stewart and Dr. Feder's paper, which was sent to Darsee's coauthors in March, 1984 by the editor of Nature, it was alleged, at p. 18, that there was an error in what was then Paper 25 (now Paper 92) regarding a purported contradiction between the text and Figure 6 concerning systolic segment lengthening. After being advised by several of the Harvard coauthors that this allegation was patently incorrect, Mr. Stewart and Dr. Feder removed the allegation from the text of the next draft. However, they did not change the numbers in the tables to reflect this deletion. When one compares the original Table 4, which listed the number of errors and discrepancies from the Harvard papers, with the present table, the number of alleged errors has not been changed for Paper 25 (now Paper 92). Because the vast majority of errors and discrepancies which Mr. Stewart and Dr. Feder have identified in the tables of their report turn out not to be errors or discrepancies at all, the accusation that the Harvard authors failed in numerous instances to correct errors in their publications cannot be credited. Furthermore, the

triviality of those few errors that Mr. Stewart and Dr. Feder have properly identified are not practices that most scientists would consider to depart from generally accepted standards.

Finally, Mr. Stewart and Dr. Feder rely heavily in their paper upon the report of the NHLBI Panel which investigated the Darsee matter. This panel of distinguished cardiovascular scientists, with a large, knowledgable staff, studied the complex matter in detail first hand over many months. They made site visits to Harvard, inspected data, and interviewed all relevant personnel. In their analysis of the laboratory and the Harvard coauthors, Mr. Stewart and Dr. Feder fail to mention the many favorable conclusions about the laboratory which this panel reached. Among these conclusions of the NHLBI Panel are the following:

> "Apart from the professional misconduct of Dr. John Darsee, the integrity and scientific capabilities of the professional and technical staff of the Cardiovascular Laboratory at the Brigham and Women's Hospital are of a high order. (page 32)
>
> "At all times . . . all officials and employees at Harvard and the Brigham and Women's Hospital cooperated fully and promptly in providing all requested information and in facilitating the panel's study in every way possible. (page 32)
>
> "(The Harvard co-authors) deserve full credit for initiating the measurement of residual radioactivity in cardiac tissues. This technique led to the important discovery that there was an absence of residual radioactivity in many of the hearts in which Dr. Darsee purported to have measured myocardial blood flow by injecting radioactive microspheres. (page 23)
>
> "The [Cardiac Research] Laboratory has had an outstanding record of productivity since its establishment . . . in 1973. It has provided research training and experience for many young investigators. (page 27)

> "In many respects, supervision in the Cardiac Research Laboratory... has been adequate and comparable to that provided in similar laboratories at other institutions." (page 28)

Furthermore, Mr. Stewart and Dr. Feder fail to mention that, one year after the NHLBI Panel's investigation, the Panel and its staff carried out a detailed site visit of the Harvard laboratory and had no fault to find.

CONCLUSION

The subject of research practices is a vital one which deserves close and continuing attention by scientists. Dishonesty and fraud strike at the very foundations of science. A better understanding of the causes and frequency of such behavior is essential to improving detection and prevention. However, science will be best served and public understanding of scientific misconduct will be advanced when the discussion of this important issue is premised on strict adherence to accuracy and on conclusions that are rigorously supported by the available evidence.

DEPARTMENT OF HEALTH & HUMAN SERVICES Public Health Service

National Institutes of Health
Bethesda, Maryland 20892

March 23, 1987

FROM: Walter W. Stewart and Ned Feder, National Institutes of Health, Bethesda, Maryland

TO: The Chairman, Science Policy Task Force of the Committee on Science and Technology, U. S. House of Representatives

SUBJECT: Response to Statement of September 18, 1986, submitted to the Science Policy Task Force by Mr. Bancroft Littlefield, Jr., Esquire

Mr. Chairman and members of the Science Policy Task Force:

We appreciate the opportunity to respond to the "Statement for the Record" submitted to the Science Policy Task Force by Mr. Littlefield. The Statement deals with a report we had previously presented to the Task Force on May 14, 1986. In this report we analyzed the professional practices of a group of 47 biomedical scientists. In particular, we tabulated some instances of professional misconduct and attempted to determine the frequency of questionable practices in this group.

Our report was drafted in 1983 and submitted to Nature. (A revised version of the report was published in the British journal Nature on January 15, 1987.) Early in 1984, Mr. John Maddox, Editor of Nature, agreed to send a draft copy of our report to 12 of the 47 scientists in the sample whose practices we were examining. Our aim was to obtain comments and criticisms that might help ensure the accuracy of the report when published.

Something then happened that was quite unusual for a scientific publication. In addition to the replies from scientists, there was a letter from a lawyer acting on behalf of a distinguished senior scientist at

Harvard Medical School who was one of those in our sample. Further letters followed from this lawyer.

It seemed likely that the lawyer's letters were intended primarily as harbingers of litigation. First, the response was not from a scientist but from a lawyer, and this alone suggested to us, as it did to journal editors, that litigation was contemplated. Furthermore, the letters contained threats of litigation that might daunt any but the most courageous or foolhardy of editors. There have been more than 20 letters and memoranda containing a total of about 175 typed pages. Excerpts follow.

— ". . . you should be aware that the case for libel, strong as it is now, would be immeasurably strengthened if you published the paper after these inaccuracies had been called to your attention." (Mr. Littlefield to Stewart and Feder. June 1, 1984.)

— ". . . please be advised that if it is your intention not to revise the paper, my review of that paper conclusively demonstrates that it is libelous and I will recommend that [my client] take appropriate legal action if it is published either in a journal or by informal circulation." (Mr. Littlefield to Stewart and Feder, with copy to Editor of Nature. July 20, 1984.)

— ". . . publication of the Stewart and Feder paper by a journal with national circulation would constitute a separate cause of action for libel in all fifty states." (Mr. Littlefield to Mr. Robert Lanman, Esquire, Office of the General Counsel, NIH. August 7, 1984.)

— "In most American jurisdictions, however, malice need not be shown to make out a claim for libel. . . . [I]n a libel action, the authors or publishers charged with defamation bear the burden of proving the truth of their allegations." (Mr. Littlefield to Stewart and Feder. November 21, 1984.)

The writer of these letters, the lawyer who so energetically represented his client, is Mr. Bancroft Littlefield, Jr. — the same person who signed the September 18 Statement for the Record submitted to the Science Policy Task Force. Nowhere in his Statement or in the covering letter does Mr. Littlefield state the capacity in which he prepared and submitted this Statement to the Task Force, and nowhere in either document does he name the distinguished senior scientist at Harvard Medical School who at one time was his client, and perhaps still is. Nonetheless, it seems

plausible that in preparing his Statement, Mr. Littlefield is continuing his role as advocate for the distinguished senior scientist — but it is of course possible that he is not. We trust that the printed record will include documentation of the fact that Mr. Littlefield previously served as attorney for one of the scientists in our sample.

In contrast to the passages from Mr. Littlefield's letters quoted above, his Statement for the Record appears on its surface to be aimed only at the factual accuracy of our report. But it is a matter of record that he has vigorously raised the possibility of a libel suit by his client if our report were published.

Whatever their purpose, Mr. Littlefield's letters had the effect of demoralizing and intimidating most editors of scientific journals considering our report for publication. At one point Mr. Littlefield wrote a journal editor that he had "specifically sought not to raise the issue of litigation in order to avoid even the appearance" that the threat of a libel action was being used to influence that journal's editorial judgment. But we can assure you, Mr. Chairman, that some journal editors were very concerned that Mr. Littlefield or the scientist who was his client or both were contemplating a lawsuit. Indeed it seems clear that this is one of the reasons they were unwilling to publish our paper.

Commenting on the growing presence of lawyers in the publication process, Dr. Benjamin Lewin, the Editor of Cell stated: "It's wildly inappropriate. It makes things much more complicated, increases the costs enormously, and exerts an inhibitory effect. But it can be quite effective, too."

Threats of litigation are almost always inappropriate, in our view, as a means of dealing with disputes over the accuracy of a scientific publication. We recommend that all scientists publicly subscribe to the idea that they will rely on public debate rather than litigation or threats of litigation in dealing with disputes over the accuracy of scientific publication.

We turn now to specific criticisms raised by Mr. Littlefield in his 25-page Statement for the Record of September 18. In the interest of brevity, we have omitted a few points we considered minor.

Item 1. Mr. Littlefield suggests on page 3 of his Statement that our report inaccurately attributes certain "motives or states of mind" to those in our sample. Mr. Littlefield has taken the quoted phrase from a letter written to us about two years ago by Mr. Maddox, the Editor of Nature. However, in the same letter Mr. Maddox also noted that we had agreed to changes that would eliminate this imputation of motives or states of mind, and the phrases in question were subsequently revised. Thus Mr. Littlefield's objection has no relevance to current versions of our report.

Item 2. In his Statement, Mr. Littlefield quotes a passage from a brief note to a commentary on our report entitled, "Do libel risks prevent publication?" which appeared in Nature (March 6, 1986). The passage was written by Mr. Maddox, and it refers to some of the scientists in the sample we analyzed:

> Strictly speaking, it is not correct that some of the un-named but identifiable co-authors threatened Nature with a libel suit . . . .

As laymen, we are impressed by the phrases in Mr. Littlefield's letters, such as:

— ". . . the case for libel, strong as it is now, would be immeasurably strengthened . . . ."

— ". . . conclusively demonstrates that it is libelous . . . ."

— ". . . I will recommend that [my client] take appropriate legal action . . . ."

— ". . . strong evidence of malice."

— ". . . a separate cause of action for libel in all fifty states."

— ". . . demonstrably libelous."

— ". . . will unlawfully libel and defame . . . ."

These are obviously threats of a libel suit.

Item 3. Mr. Littlefield indicates in his Statement that journal editors reacted unfavorably to our report. As examples, he quotes the remarks of three journal editors. But his chosen quotations are irrelevant, or misleadingly selective, or quoted out of context. Not one of the three editors criticizes the accuracy of our facts or analysis. One editor, Dr. Thomas P. Stossel, in the passage quoted by Mr. Littlefield, writes:

> — "I found the paper to be overstated, woolly, and excessively judgemental and moralistic."

But in the same letter, in a passage not quoted by Mr. Littlefield, this editor commented on "the (to me) rather chilling emanations from the lawyers" and went on to say:

— ". . . the paper does contain some disturbing and in my opinion unassailable points. I am further disturbed by the trail of events that this paper has elicited, namely to be strung along by several prominent journals and finally shut off because of the legal pressures brought to bear. . . . I am terribly disturbed about publication practices being directed by attorneys." (Thomas P. Stossel, Editor of Journal of Clinical Investigation to Arnold Relman, Editor of New England Journal of Medicine. August 30, 1985.)

Furthermore, Mr. Littlefield fails to quote other editors whose comments were favorable:

— "I think that what you have done is an important public service . . . ." (John Maddox, Editor of Nature. October 21, 1983.)

— "The consensus of opinion has been that it would be overwhelmingly in the public interest for this article to be published." (Benjamin Lewin, Editor of Cell. September 4, 1985.)

— ". . . I am sorry to write that we see no feasible way to use it in American Scientist. Both our publisher and our long-time legal advisor believe that the legal costs incurred would be prohibitive. I am disappointed that we have to send you this negative decision, and that there is nothing we can do to help. It has been sobering to look at your and Dr. Feder's clear and convincing report." (Michelle Press, Editor of American Scientist. September 6, 1985.)

Item 4. In his Statement, Mr. Littlefield claims that reviewers and editors have found our report to be inaccurate. This claim is simply false. Certain readers have

— criticized our sample as nonrepresentative,

— objected to the fact that our examination of the practices of the 47 coauthors is limited to their research with Dr. John Darsee,

— said that the scientific community would be damaged by our report,

— made other criticisms.

But almost without exception, readers have not criticized the accuracy of our analysis of the coauthors' practices. Indeed readers, including experts in cardiology, have commented favorably on the wealth of documentation in our report and on the accuracy of our analysis.

Item 5. In "Criticism No. 1" of his Statement, Mr. Littlefield comments unfavorably on our analysis of misleading statements. He presents a series of subsidiary and in some cases irrelevant objections to this part of our report. We discuss these objections, all of which we believe are without merit, in the Appendix, part A.

Item 6. In his "Criticism No. 2," Mr. Littlefield states that two sentences in a paper published from Harvard "clearly and explicitly" refer the reader to a previous Harvard paper containing much of the same data.

Mr. Littlefield's assertion is irrelevant to our point. In our report we stated that the two citations of the first paper provide no hint that some of the same data are being republished and that some of the same animals were used. Thus readers would be misled about the true relationship between the papers. Mr. Littlefield is silent on this point.

Dr. Lewin, the Editor of Cell, agreed with our conclusions. Previously, over a period of nine months, he had reviewed our paper for publication, and he had considered the objection raised on this point in letters from Mr. Littlefield to Cell — the very same objection now restated as "Criticism No. 2." Dr. Lewin was quoted in a New York Times article (April 22, 1986) as saying that the second paper cites the first "in a way that will clearly mislead the reader on their relationship." Our report has been examined by many referees and other outside readers, and to our knowledge none has objected to our presentation of the facts or our conclusion on this issue.

In our report we reproduce in its entirety the paragraph in question just as it was published by the Harvard authors. A reader can therefore examine this paragraph and decide for himself if our conclusion is correct.

Item 7. We indicate in our report that abstracts republished by some of those in our sample have been titled misleadingly. In his Statement, Mr. Littlefield offers an explanation: he writes of changes intended "to enhance its quality and clarity," and he refers to "tailoring or improving the abstracts." But this does not deal with the phenomenon of markedly altered titles over nearly identical texts.

At the end of the Appendix, we have reproduced the four pairs of abstracts that we cite in our report. We think someone seeing only the

titles of each pair would be misled and would conclude, incorrectly, that the two abstracts were different. In his Statement, Mr. Littlefield fails to deal with this crucial point.

Item 8. In his "Criticism No. 3," Mr. Littlefield complains that we are withholding information that would enable a reader to check our report.

As we have pointed out to Mr. Littlefield in correspondence, our report satisfies to an unusual degree the recommendation that original data on which a research report is based should be available to others. All our original data are readily available: our analysis is based on journal articles, abstracts, and committee reports, all of which are public documents. Thus from the outset, in 1984, the original data have been available to Mr. Littlefield or any other interested party. Moreover, about two years ago we wrote Mr. Littlefield offering to help him obtain photocopies of the journal articles, abstracts, or committee reports.

In his Statement, Mr. Littlefield has also asked us to "discuss and review" with him the details of our reasoning and our analysis of the original data. We declined to do this during the period before our report was published. We are not aware of any tradition that a scientist is under an obligation to provide the details of his reasoning and analysis, before publication, to anyone (except of course his coauthors and those directly involved in publication). In particular, there would appear to be no obligation to provide these details to a lawyer using threats of litigation as a means of preventing publication.

We had every reason to believe that any material we gave Mr. Littlefield would be used in further attempts to block publication. Now that our report has been published (January 15, 1987), we shall make the details of our analysis available on request to any interested party.

Item 9. Mr. Littlefield asserts in his Statement that a "careful review" shows that we were almost entirely incorrect in our identification of 15 errors in papers coauthored by his client.

In the Appendix, part B, we discuss in detail Mr. Littlefield's objections, which we think are unsupported by the facts. Furthermore, we think his statement that "most scientists" would dispute our interpretation

is simply wrong. As we have mentioned (in Items 3 and 4) and documented previously, most scientists who have spoken to us or written to us or journal editors have commented favorably on our analysis.

In his Statement for the Record, Mr. Littlefield fails to identify the scientists, lawyers, or others who carried out the "careful review" showing that our identification of errors is incorrect.

Many of the papers in question contain defects discoverable almost at a glance. In fact, Mr. Littlefield's client has vouched for this himself. About four years ago he looked at some of published papers coauthored by Darsee on work at another medical school. He indicated in a memo to NIH that some of the published measurements were unbelievable. According to an article in Science (April 1, 1983), on reviewing the papers he became concerned "almost immediately" about the integrity of three of them.

Item 10. Mr. Littlefield asserts (page 23) that we corrected an error in the text in an early version of our report, but failed to make a corresponding correction in the table. Mr. Littlefield is right about the error in the text. In the spring of 1984 we deleted from the text of our report an assertion that ref. 92 contained a particular error; we did this on the grounds that the assertion was at least doubtful and might in fact be wrong. At about the same time one of the Harvard coauthors independently discovered this questionable passage. In the same time period we discovered another error in ref. 92, and for this reason, the total number of errors we counted for ref. 92 and recorded in the tables was unchanged. Thus Mr. Littlefield is wrong in suggesting that the numbers in the tables do not reflect the errors discussed in the text. The version of our tables previous to the spring of 1984 accurately reflected the text, and the same is true of subsequent versions, including the present version.

Item 11. In his Statement, Mr. Littlefield devotes a full page to the fact that a committee (the NHLBI Panel) investigating the Darsee affair had some words of praise for his client's research group and also that the committee "had no fault to find" when it revisited the laboratory a year after its initial investigation. All this is of course entirely irrelevant to the accuracy of our report.

Incidentally, Mr. Littlefield fails to note that the NHLBI panel was critical of some of his client's actions. For example, the panel's report

referred to the investigation of Darsee's research conducted privately inside the Harvard laboratory by the two senior scientists as "insufficiently rigorous," and the report criticized the failure to include outside scientists as investigators.

APPENDIX

Part A. See Item 5 above.

In his "Criticism No. 1," Mr. Littlefield challenges the main conclusion of one section of our report. His criticisms are mainly peripheral or irrelevant. In this section of our report, we state that the two senior coauthors at Harvard published statements that they knew or should have known were inaccurate. Our evidence comes from the NHLBI panel report, which investigated the Darsee affair, and from communications written to the NIH by the two senior Harvard coauthors. The evidence in these public documents is confirmed by Mr. Littlefield's Statement for the Record.

As we discussed in our report and now summarize here, the question centers around the use of data from "historical" control dogs as distinct from contemporaneous control dogs. In the five Harvard papers we analyzed in this part of our report, the experimenter made use of the records of previously obtained control data (historical controls) instead of making observations and obtaining new control data from a living dog. The use of historical controls was sanctioned by the two senior Harvard coauthors, as was the failure to describe that use on publication.

There is good evidence, based on statements made by both Mr. Littlefield and his former client, that the latter sanctioned the publication of a description that, in fact, was inaccurate. This is the same conclusion we reached in the analysis presented in our report. Moreover, in a letter to the Editor of Cell (October 22, 1985), Mr. Littlefield himself suggested the following sentence for use in our report:

> It would appear from their subsequent statements that the two coauthors of the five papers knew of and authorized the limited use of historical controls in the five papers (although certainly not the extent to which or the manner in which Darsee actually used those controls), as well as the failure to describe that use on publication.

This statement by Mr. Littlefield, and others by his client that we quote in our report, indicate that the latter was aware at the time of publication that the published statement was inaccurate. In our report, however, we make a weaker assertion: that he "knew or should have known" of the inaccuracy at the time of publication.

Mr. Littlefield misstates a major assertion of this part of our report. He claims (page 8) that in our report we say that a certain published description was incomplete — he calls it a "failure adequately to describe" — in the group of five papers of which his client was a coauthor. Mr. Littlefield's characterization is wrong. In our report we indicate that statements published in the papers were false, not incomplete. False statements are obviously different from incomplete statements, and generally more serious.

Here are some criticisms made by Mr. Littlefield in this part of his Statement.

1. In some parts of his Statement, Mr. Littlefield implies (though he does not state) that the five Harvard papers criticized in our report give an accurate description of experimental methods. In fact, the papers contain untrue statements about the method of selecting the controls. For example:

> Twenty-five minutes after occlusion of the LAD [in the heart of an anesthetized dog], the remaining dogs were randomized to treated or control group on the basis of a random odd or even number generated by a computer program. [Quoted from Circulation, 1981, 63, 29-35.]

Here is why the above published statement describing the assignment of dogs to treated and control groups is untrue. According to the quoted description, each dog was assigned to the experimental or control group partway through the surgical procedure — after the dog was anesthetized, the chest was opened, and a coronary artery (LAD) was occluded. But if this were true, then it would be impossible to introduce previously recorded historical control data into the experiment. The actual use of historical control data — a fact withheld from readers — is incompatible with the published statement describing control data obtained contemporaneously from living dogs after a specific randomization procedure. Therefore, the published description of randomization based on a computer program is necessarily untrue. The description is not incomplete, but false. The same applies to other papers in this group of five papers.

2. Mr. Littlefield indicates that the method of selecting control dogs was not "critical to another investigator's ability to assess or repeat the experiment." We believe that Mr. Littlefield is wrong and that many scientists would dispute Mr. Littlefield's claim that the use of historical versus contemporaneous controls in experiments of this sort is unimportant and not worth mentioning. When the information is withheld, the reader, unaware of the use of historical controls, is denied the possibility of judging the appropriateness, from a scientific viewpoint, of the decision by the senior Harvard scientists to use historical controls in this particular set of experiments. Certainly readers should not have been presented with an inaccurate description of the experimental procedure.

The published statements that were inaccurate had the effect of making the data and the conclusions look better to the reader than they actually were.

Mr. Littlefield defends the use of historical controls as a means of "minimizing the unnecessary sacrifice of dogs." But for this very reason — the need to minimize the sacrifice of dogs — experiments should be done carefully and reported truthfully.

Clearly the publication of a false description of an experiment compromises the accuracy and value of the scientific literature.

3. On page 7 of his Statement, Mr. Littlefield describes as "similar" the experimental conditions for the contemporaneous control dogs and the historical controls.

— In our report we list in a table the many dissimilarities that actually existed. We think that most scientists would dispute the claim that the experimental conditions were "similar."

— Readers were not told that dogs from other experiments were being used to provide data. Thus the information was withheld from readers showing that in any particular paper there were differences in the experimental conditions to which the dogs were subjected. Certainly there was no way readers could judge for themselves whether the dogs were being exposed to conditions that were dissimilar or, as Mr. Littlefield asserts, "similar."

4. Mr. Littlefield advances two further arguments to justify his client's actions. Both are irrelevant to the accuracy of our report.

— He suggests that the experiments did not really involve historical controls. His reason: all the experiments were done "within a relatively short period of time of each other." The essential point, however, is not the duration of the time period, but rather that the experimental procedure was described inaccurately in the five Harvard papers.

— Mr. Littlefield places great emphasis on Darsee's failure to comply with lab policy. For our analysis, Darsee's behavior is irrelevant to the question we raised in our report: was the experimental procedure (as approved by the senior authors) described correctly or incorrectly on publication? Darsee's undoubted misbehavior has no bearing on the fact that the senior authors approved the publication of an inaccurate and misleading description of an experimental procedure that they themselves had planned.

Part B. See Item 9.

Mr. Littlefield asserts in his Statement that we were almost entirely incorrect in our identifications of 15 errors in his client's published papers. Mr. Littlefield's specific objections follow.

1a.

Mr. Littlefield gives an instance of a supposed inaccuracy in our report, in our analysis of Figure 9 in one of the Harvard papers (dealing with a certain measure of cardiac function). He criticizes as being "without merit" our conclusion that Figure 9 actually depicted a large improvement, rather than a small improvement as asserted by the authors in their published paper.

We based our conclusion (i) on our own analysis, (ii) on consultation with expert cardiologists familiar with precisely this kind of experiment and its interpretation, and (iii) on the apparent confirmation of our conclusion by one of the two senior Harvard coauthors: "Figure 9 - subepicardium shows that recovery is only approximately 50% of baseline value." (This is in a letter from this coauthor to Mr. Maddox, April 12, 1984).

The improvement is described explicitly in our report as a 50% improvement. This allows readers to judge whether or not we are correct in describing this as a "large improvement."

1b and 1c.

In his Statement Mr. Littlefield suggests we are demanding unrealistically high standards in the graphs and drawings for these papers

and states that the discrepancies between the figures and the numerical values are acceptable because they fall "within the limitations of the artwork."

On the contrary, we think most scientists would agree that graphs in scientific publications are generally intended to convey quantitative information accurately.

In his Statement, Mr. Littlefield criticizes our analysis of a figure depicting ultrasonic crystals implanted in a heart. The figure could easily have been drawn to show a distinct separation between ultrasonic crystal and necrotic tissue, as described in the text. What the figure actually showed was contact between crystal and tissue. This error should have been corrected before publication.

Similarly, Mr. Littlefield criticizes our assertion of a particular discrepancy between the text and a graph. We believe the length of the error bar drawn in a figure (i.e., 3%), in comparison with the numerical value (1.1%) given in the text, is outside the limits of appropriate accuracy for scientific artwork. The discrepancy between the two values for the error bar is three-fold, it is obvious, and it would have been easy to correct.

An example of another similar error is in Fig. 3 (percent midmyocardial segment shortening of control dogs). Here both the position of the mean and the length of an error bar are drawn inaccurately (-16.7±3.5% in the text vs. -11.9±2.2% in the figure).

1d.

Mr. Littlefield refers to two graphs in our reference 92 and states that it is "untenable to claim that the similarity in the error bars was so pronounced as to be inherently suspicious."

At the end of the Appendix we reproduce the graphs in question. We think that an inspection of these graphs with the care that one would expect of a coauthor should immediately have raised a serious question about the similarity of the error bars, especially if the coauthor were a scientist with considerable experience in research. Once such questions were raised, the next step, of course, would have been a reassessment of the original data, which presumably would have disclosed a serious problem.

Mr. Littlefield points out that his client was not alone in failing to recognize the implausible similarity of error bars. In defense of his client, Mr. Littlefield points out that the peer-reviewing system (i.e., the reviewers chosen by the journal editor) also failed to raise questions about these error bars. The failure of the peer-reviewing system in no way excuses the authors. The primary reponsibility for the accuracy of research and publication clearly lies not with the journal reviewers, but at the source, with the authors themselves.

2a.

Mr. Littlefield objects to a passage in our report on the grounds that a discrepancy we discuss is very minor. We agree that the discrepancy is very minor. Mr. Littlefield fails to note that in our report we emphasize that we present this discrepancy "as an example of an error of very minor significance." He also fails to note that in our report, we quote the published material in question so that a reader can make up his own mind.

2b.

Mr. Littlefield criticizes our conclusion about two slightly different numbers for what seems to be the same measurement (of ATP concentrations). We counted this as a discrepancy. Mr. Littlefield does not say our conclusion is wrong, but he gives a hypothetical explanation: the numbers might have been calculated in two different ways, and this hypothesis, he argues, could explain the difference. Previously (for example, on July 20, 1984), Mr. Littlefield defended this particular pair of discrepant values on different grounds: "At most, this is a minor error." No mention was made of differing methods of calculation.

We think that a paper prepared with adequate care should not contain a discrepancy whose explanation is not apparent to the reader.

2c.

Mr. Littlefield gives another instance of a supposed inaccuracy in our report (involving dp/dt values in different publications). He indicates that our analysis "is based on hindsight" and that it is unreasonable to expect that it would have been noticed at the time.

We believe that simple, careful inspection — not a recalculation, as Mr. Littlefield asserts — by an experienced senior scientist should lead the scientist to question the consistency of the data.

Mr. Littlefield suggests that, in examining papers for errors, we are using standards which are unreasonably high and which would be appropriate for authors of a paper only if there were a suspicion of fraud. On the contrary, this kind of checking before publication is in our view a part of normal scientific procedure. It does not indicate or require prior doubts about the integrity of one's coworkers. There is only the recognition that it is easy for anyone to make mistakes, coupled with the belief that occurrences of these mistakes should be kept to a minimum in a published research report.

We doubt that most scientists would accept as valid the arguments given by Mr. Littlefield in his Statement for the Record.

DIASTOLIC PRESSURE-VOLUME CHARACTERISTICS DURING POSITIVE END EXPIRATORY PRESSURE VENTILATION: PARADOXICAL RISE IN LEFT VENTRICULAR FILLING PRESSURE.

J.R. Darsee, J.R. Mikolich,* P.F. Walter,* R.C. Schlant,**
Department of Medicine, Emory University Hospitals, Atlanta, Georgia

Controversy exists whether positive end expiratory pressure (PEEP) ventilation lowers cardiac output by reducing left ventricular (LV) preload, increasing LV afterload, or by combination effects. Accordingly, 12 anesthetized mongrel dogs were studied both closed and open chest during 0,5,10,15 and 20 cm H_2O PEEP with high-fidelity LV systolic pressure, left (LVEDP) and right ventricular end diastolic pressure (RVEDP), and dP/dt, cuff or catheter aortic flow probe, sonomicrometry of LV dimensions, with a right atrial reservoir for rapid changes in RVEDP. With chest closed, changes were greatest with 20 cm H_2O PEEP: LVEDP ↑ from 4.1±1.2 to 8.2 2.4 mmHg as LV systolic fell from 124±8 to 98±4 mmHg, with no change in heart rate, a ↑ in stroke volume from 28±3 to 15±3 ml, ↑ in LVED dimension from 30.2±2 to 27.4±1.4 mm, ↓ in peak wall tension from 110±8 to 64±6 g/cm^2, and ↓ in maximum -dP/dt from -1800±120 to -790±80 mmHg/sec, while RVEDP ↑ from 3.9±0.8 to 10.3±2.1 mmHg (for all changes p<.01). With lung and pleural factors eliminated with chest open, LVEDP rise was less (to 6.9±0.8 mmHg) than with chest closed. With a sudden drop in RVEDP (by reservoir) to pre-PEEP values for 2 beats (not affect LV filling volume for ≥5 beats), LVEDP fell to <2 mmHg in all dogs. We conclude: 1) a paradoxical ↑ in LVEDP with ↑ LVED dimension (↓ LV compliance) with PEEP is caused primarily by elevated RVEDP transmitted across the septum; and 2) with RV factors transiently removed, PEEP is associated with ↓ LVEDP and ↑ LVED dimension as would be expected.

Clin. Res. 27, 726A (1979).

PARADOXICAL RISE IN LEFT VENTRICULAR FILLING PRESSURE WITH POSITIVE END EXPIRATORY PRESSURE VENTILATION: A REVERSED BERNHEIM EFFECT.

John R. Darsee, MD, J. Ronald Mikolich, MD, Paul F. Walter, MD, Robert C. Schlant, MD, FACC. Emory University School of Medicine, Atlanta, Georgia

Controversy exists whether positive and expiratory pressure (PEEP) ventilation lowers cardiac output by reducing left ventricular (LV) preload, increasing LV afterload, by influencing contractility, or by combination effects. Accordingly, 12 anesthetized mongrel dogs were studied both closed and open chest during 0,5,10,15 and 20 cm H_2O PEEP with high-fidelity measurements of LV systolic pressure, LV end-diastolic pressure (LVEDP), and dP/dt, cuff or catheter aortic flow probe, and sonomicrometry of LV dimensions. The right atrium was connected to a reservoir for rapid changes in right ventricular (RV) EDP measured with a high-fidelity catheter. With chest closed, changes were greatest during 20 cm H_2O PEEP: LVEDP rose from 4.1±1.2 to 8.2±2.4 mmHg as LV systolic fell from 124±8 to 98±4 mmHg, with no change in heart rate, a ↑ in stroke volume from 28±3 to 15±3 ml, ↑ in LVED dimension from 30.2±2 to 27.4±1.4 mm, ↓ in maximum negative dP/dt from -1800±120 to -790±80 mmHg/sec, and ↓ in peak wall tension from 110±8 to 64±6 g/cm^2, while RVEDP rose from 3.9±0.8 to 10.5±2.1 mmHg (for all changes p<.01). With lung and pleural factors eliminated in open chest dogs during 20 cm H_2O PEEP, LVEDP rise was less (to 6.9± 0.8 mmHg) than with chest closed. With a sudden drop in RVEDP (using the reservoir) to pre-PEEP values for 2 beats (which would not affect LV filling volume for ≥ 5 beats), LVEDP fell to < 2 mmHg in all dogs. We conclude: 1) a paradoxical ↑ in LVEDP with ↓ LVED dimension (↓ LV compliance) during PEEP is caused primarily by elevated RVEDP transferred across the septum; and 2) with RV factors transiently removed, PEEP is associated with ↓ LVEDP and ↑ LV wall motion.

Am. J. Cardiol. 45, 476 (1980).

INTERVENTION PRODUCING EARLY RECOVERY OF SEGMENTAL FUNCTION IN SALVAGED ISCHEMIC MYOCARDIUM FOLLOWING EXPERIMENTAL CORONARY OCCLUSION.

J.R. Darsee, R.A. Kloner,* E. Braunwald,**
Harvard Medical School and Peter Bent Brigham Hospital, Boston, Massachusetts

Although agents have been shown expertimentally to protect ischemic myocardium, a critical unanswered question is whether function is preserved in the salvaged tissue. Accordingly, 28 open-chest dogs had measurements of percent segment length shortening (ISS) and peak velocity of shortening (PVSS) by sonomicrometry of equatorial midmyocardial ischemic segments (IS). Fifteen minutes following left anterior descending coronary occlusion, microspheres were injected for measurement of regional myocardial blood flow (RMBF) and dogs were treated with flurbiprofen, an anti-inflammatory drug (1mg/kg IV; n=13) or saline (n=15). Following, 60 min occlusion, ISS and PVSS were followed during 5 hrs of reperfusion. Myocardium at risk was determined by in vivo injection of methylene blue which left unperfused areas unstained and necrosis by triphenyl tetrazolium chloride and planimetry of 5mm transverse sections. RMBF within the measured IS was 0.36±.02 and 0.34±.03 ml/min/g in treated (T) and control (C) dogs, respectively (p=NS). In C dogs, 18±3% of the myocardium at risk became necrotic compared to 0% necrosis in T dogs (p<.01). In C dogs, passive lengthening occurred during 5 hrs of reperfusion; in T dogs, ISS returned to preocclusion values by 5 min reperfusion although PVSS remained depressed (preocclusion=18±2 mm/sec; 5 hrs reperfusion=11±2 mm/sec; p<.01).

In conclusion, flurbiprofen not only protects ischemic myocardium from necrosis but allows recovery of function unlike equally ischemic, but untreated myocardium. Earlier functional recovery following prolonged ischemia indicates an important role for intervention therapy.

Clin. Res. 28, 163A (1980).

PRESERVATION OF ISCHEMIC SEGMENT FUNCTION IN NON-NECROTIC MYOCARDIUM BY A NEW ANTIINFLAMMATORY AGENT 913

John R. Darsee, Robert A. Kloner, Eugene Braunwald,
Harvard Med. School and Peter Bent Brigham Hospital, Boston, Massachusetts

Although agents have been shown to protect ischemic myocardium, a critical unanswered question is whether function is preserved in salvaged tissue. Accordingly, 28 open-chest dogs had measurements of segment length (SL) and dSL/dt in equatorial midmyocardial ischemic segments. 15 min after coronary occlusion, regional blood flow (RMBF) was measured with microspheres and dogs were treated with either flurbiprofen, an antiinflammatory agent (1mg/kg IV; n=13) or saline (n=15). Following 60 min of occlusion, SL and dSL/dt were recorded for 3hrs reperfusion. Hearts were sliced transversely and necrosis quantified by triphenyltetrazolium chloride. RMBF within the evaluated ischemic SL was .34±.02 and .34±.03 ml/min/g in treated (T) and control (C) dogs, respectively (p= NS). In C dogs 6±1% of the left ventricle (all subendocardial) became necrotic; no necrosis occurred in T dogs. In C dogs, SL passive lengthening (-2.1±3mm) persisted during 3hrs reperfusion; whereas SL shortening returned to preocclusion values (+1.9± 2mm) in T dogs within 3min reperfusion (p <.01), although peak dSL/dt remained depressed (preocclusion=18±2mm/sec; 3 hrs reperfusion=11±2mm/sec, p<.05).

In conclusion, flurbiprofen not only protects ischemic myocardium from necrosis but preserves function in non-necrotic ischemic tissue which does not occur with reperfusion alone.

Circulation 62, III-239 (1980).

● PROLONGED METABOLIC, FUNCTIONAL, AND ULTRASTRUCTURAL ABNORMALITIES FOLLOWING TRANSIENT MYOCARDIAL ISCHEMIA WITHOUT INFARCTION.

R.A. Kloner, L.W.V. DeBoer*, J.R. Darsee, J.S. Ingwall*, and E. Braunwald**. Harvard Medical School/Peter Bent Brigham Hospital, Boston, Massachusetts

Brief 15 minute periods of temporary coronary occlusion (CO) followed by reperfusion are not associated with tissue necrosis, but the biochemistry, ultrastructure, and function of this injured myocardium, after several days of reperfusion has not been studied in detail. Therefore, 14 anesthetized, open chest dogs received 15 minutes of left anterior descending (LAD) CO followed by 72 hours of closed chest reperfusion. Myocardial biopsies were obtained from the LAD and circumflex (CX) beds at 15 min and 72 hours for ATP (nmoles/mg protein) levels, and electron microscopy (EM). At 72 hours, tissue was also obtained from the LAD and CX beds for routine histology. Segment function in the LAD bed was measured with ultrasonic crystals. 15 min after CO ATP was 38±3 in the CX bed and 20±4 in the LAD bed ($p<.001$). With 72 hours of reperfusion ATP in the CX bed remained normal but in the LAD bed remained depressed (29±3, $p<.025$). At 15 min of CO, LAD bed segment shortening (SS) was replaced by passive lengthening; at 72 hrs of reperfusion, SS still remained reduced at 57±5% ($p<.03$) of preocclusion value. At 15 min of CO glycogen loss and I bands were present; at 72 hours there were areas of localized sarcoplasmic reticular and intermyofibrilar edema and glycogen loss. At 72 hours no necrosis was present in either LAD or CX beds by both routine histology or electronmicroscopy. Therefore, brief transient episodes of myocardial ischemia which do not result in myocardial infarction may still result in metabolic, functional, and ultrastructural abnormalities which persist for at least 72 hours.

Clin. Res. 28, 615A (1980).

PERSISTENT MYOCARDIAL ABNORMALITIES FOLLOWING BRIEF PERIODS OF TEMPORARY CORONARY OCCLUSION NOT ASSOCIATED WITH NECROSIS. 292

R.A. Kloner, L.W.V. DeBoer, J.R. Darsee, J.I. Ingwall and E. Braunwald. Harvard Medical School and Peter B. Brigham Hospital, Boston, MA

Although brief 15 minute periods of temporary coronary occlusion (CO) followed by reperfusion (R) are not associated with tissue necrosis, the biochemistry, ultrastructure, and function of this injured myocardium, after several days of R has not been studied in detail. Therefore, 14 anesthetised, open chest dogs received 15 minutes of left anterior descending (LAD) CO followed by 72 hours of closed chest R. Myocardial biopsies were obtained from the LAD and circumflex (CX) beds at 15 min and 72 hours for ATP (nmoles/mg protein) levels, and electron microscopy (EM). Segment function in the LAD bed was measured with ultrasonic crystals. 15 min after CO ATP was 38±3 in the CX bed and 20±4 in the LAD bed ($p<.001$). With 72 hours of R ATP in the CX bed remained normal but in the LAD bed remained depressed (29±3, $p<.025$). At 15 min of CO, LAD bed segment shortening (SS) was replaced by passive lengthening; at 72 hrs of R, SS still remained reduced at 57±5% ($p<.03$) of preocclusion value. At 15 min of CO glycogen loss and I bands were present; at 72 hours there were areas of localized sarcoplasmic reticular and intermyofibrilar edema and glycogen loss - but no necrosis. Therefore, brief periods of temporary CO not associated with necrosis may result in myocardial abnormalities 3 days after reperfusion.

Circulation 62, III-80 (1980).

● EARLY PATHOLOGIC DETECTION OF ACUTE MYOCARDIAL INFARCTION.

R.A. Kloner, J.R. Darsee, L.W.V. DeBoer*, N. Carlson* (intr. by E. Braunwald**) Department of Medicine, Harvard Med. School/Peter B. Brigham Hospital, Boston, MA

The use of tetrazolium stains for identifying myocardial infarction during the early stages of ischemic injury has become increasingly popular in both experimental and clinical studies. However, the accuracy of these stains in delineating infarctions prior to the onset of well defined histologic necrosis is unknown. Therefore 8 open-chest, anesthetized dogs were subjected to high left anterior descending coronary artery occlusions for 6 hours. The hearts were then excised and "breadloafed" into 3mm thick transverse slices which were incubated in triphenyltetrazolium chloride (TTC) for 10 minutes. Three to 5 tissue samples were obtained from each of three areas of two transverse ventricular slices: from TTC unstained areas (infarction) and TTC stained areas both close to the infarct and remote from the infarct for histologic and electronmicroscopic analysis. All 8 dogs had well defined areas of unstained tissue which appeared pale grey next to the brick red TTC stained tissue. Routine histology of TTC unstained areas revealed interstitial edema, waviness of myocytes, and occasional polymorphonuclear leukocytes but no definite necrosis; TTC stained areas both close to and remote from the infarct appeared normal by histology. Ultrastructure of the TTC stained tissue both close to and remote from the infarct showed only modest loss of glycogen and relaxation of myofibers. Ultrastructure of the TTC unstained tissue in all 8 dogs showed numerous mitochondrial amorphous dense bodies, intracellular edema, nuclear chromatin clumping and margination, sarcolemmal disruption, and interstitial edema. Therefore TTC is a reliable stain for localizing myocardial cells which ultrastructurally appear irreversibly injured by ischemia, prior to the development of well-defined necrosis as assessed by routine histology.

Clin. Res. 28, 614A (1980).

CAN TETRAZOLIUM STAINS RELIABLY IDENTIFY MYOCARDIAL INFARCTION PRIOR TO DEFINITE HISTOLOGIC NECROSIS? 939

R.A. Kloner, J.R. Darsee, L.W.V. DeBoer, N. Carlson, Harvard Medical School and Peter B Brigham Hospital Boston, Massachusetts

The use of tetrazolium stains for identifying myocardial infarction during the early stages of ischemic injury has become increasingly popular in both experimental and clinical studies. However, the accuracy of these stains in delineating infarctions prior to the onset of histologic necrosis is unknown. Therefore, 8 open-chest anesthetized dogs received high left anterior descending coronary artery occlusion for 6 hours; the hearts were excised, "breadloafed" into 3mm thick transverse slices and incubated for 10 minutes in triphenyl tetrazolium chloride (TTC). Tissue was sampled from TTC unstained areas and TTC stained areas close to and remote from the unstained area for electron microscopy (E.M.) and routine histology using hematoxylin and eosin staining. E.M. of TTC stained regions showed only mild glycogen loss and myofiber relaxation. E.M. of TTC unstained tissue in all dogs showed numerous amorphous mitochondrial densities, intracellular edema, sarcolemmal breaks, and nuclear alterations. By routine histology TTC stained areas appeared normal while TTC unstained tissue showed interstitial edema and some areas of wavy fibers, but no definite necrosis. Therefore, TTC is a reliable stain for localizing myocardial cells which ultrastructurally appear irreversibly injured by ischemia, prior to the development of well-defined histologic necrosis.

Circulation 62, III-246 (1980).

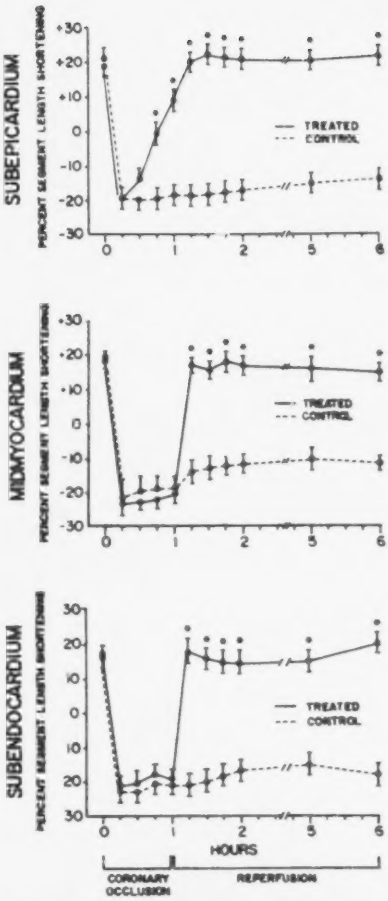


FIGURE 3 Separate graphs showing %SS in the subepicardium, midmyocardium, and subendocardium of control and treated dogs during 1 h of coronary occlusion and 5 h of reperfusion. Active systolic shortening values are represented by positive numbers on the ordinate and systolic passive lengthening values by negative numbers. Regional function was different between control and treated dogs only during reperfusion in the previously ischemic subendocardial and midmyocardial segments, but improvement in function began to occur in the subepicardium of treated dogs even while the coronary artery was still occluded, and became significantly different from control within 30 min of the time of drug administration. Values shown are mean ± SEM for each time period. *, statistically different from respective value in control dogs ($P < 0.01$).

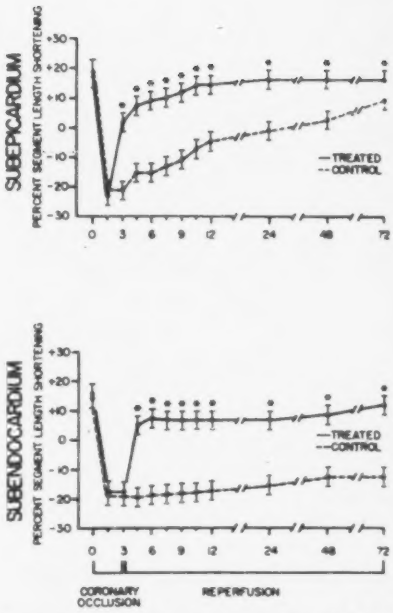


FIGURE 9 %SS in the subepicardium and subendocardium of control and treated dogs during 3 h of coronary occlusion followed by 72 h of reperfusion. Values shown are mean ± SEM for each time period. *, statistically different from respective value in control dogs ($P < 0.01$).

Figures reproduced from:
J. Clin. Invest. 68, 225-239 (1981).

APPENDIX 2

ARTICLE BY W. STEWART AND N. FEDER AND COMMENT BY E. BRAUNWALD IN NATURE MAGAZINE

NATURE VOL. 325 15 JANUARY 1987

The integrity of the scientific literature

Walter W. Stewart and Ned Feder

A case of admitted scientific fraud has shed new light on the system that ensures the integrity of the scientific literature. Certain lapses from generally accepted standards of research may be more frequent than is commonly believed.

We wish to report the results of an unusual investigation. In May 1981, the colleagues of a respected young scientist were shocked to discover that he was a forger of data. Initially, Dr John Darsee confessed to a single forgery, but it was subsequently found by his colleagues and by three investigating committees that he had fabricated much of the data that formed the basis of his more than 100 publications[1-139] over a period of about three years.

Although many of these publications embody data now acknowledged to be fraudulent, we have scrutinized these publications for internal consistency, accuracy and completeness. Our sample consists of the 109 publications by Darsee with the 47 scientists who were his co-authors, all of whom worked at two leading US medical schools. We also studied the reports of the committees set up to investigate the Darsee affair. What is special about the publications in our sample is that, as a by-product of the extensive investigations into Darsee's work, the circumstances surrounding their publication were known in unusual detail.

Our objective was to discover whether the published scientific reports would throw light on the vigilance of referees, of editors of journals and of Darsee's co-authors in meeting the standards conventionally accepted as necessary in the scientific literature. We have found many of the reports to be flawed in ways that could have been recognized by those concerned with their publication.

We emphasize the distinction between the original forgery and the lapses from standard publication practices we have found in our "sample of convenience"[140]. Other studies will be needed to determine whether or not our results are representative of clinical scientists, of biomedical scientists or of scientists in general.

Background

The circumstances of the Darsee affair have been widely reported[131-140]. In 1978 to 1981 inclusive, he was author or co-author of 18 full-length research papers published in major biomedical journals and of about 100 abstracts, book chapters, letters, reviews and short papers in clinical and experimental cardiology[1-139]. In part because he was so prolific, he was highly regarded first at Emory University School of Medicine and then at Harvard Medical School.

Five months after Darsee's confession, in May 1981, to a single act of data fabrication, it became apparent that there had been more than one such episode. Ultimately, three committees were appointed to investigate. Harvard appointed a committee headed by R.S. Ross, which issued its report[131] in January 1982. The NHLBI (National Heart, Lung, and Blood Institute) appointed a panel headed by H.E. Morgan to investigate the circumstances at Harvard, where most of the research had been supported by the National Institutes of Health (NIH); this report[138] was completed in June 1982. Finally, at Emory, an investigating committee headed by N.C. Moran was formed, and issued its report[132,139] in March 1983.

Shortly after the appearance of the report of the Moran committee, and almost two years after the frauds themselves had come to light, we embarked on our study. We were especially careful to discriminate between the direct effects of Darsee's forgeries, which are not the subject of this report, and the defects in the papers of which he was a co-author. We recognize that some may disagree with some of our decisions as to what constitutes a defect, and with our judgements of the degree to which Darsee's co-authors may be held to have been responsible.

Darsee's publications were in experimental and clinical cardiology, a field of research quite different from our own. For this reason, we have no comments to make on the pointedness or importance of the research described in the publications or about the design of the experiments. No doubt we may also have overlooked some errors in the publications because of our lack of first-hand knowledge of the field. At the same time, we believe that our position as outsiders had the advantage that none of those who published with Darsee was known to us personally or professionally. It may also be an advantage that we have been able to carry out our analysis in the same way that other scientists outside the field of cardiology might approach such a task.

Of Darsee's 47 co-authors, 24 worked at Emory University School of Medicine and 23 at Harvard Medical School. Biographical data were found for 44 of the 47 co-authors. Thirty-nine were apparently MDs, one was a PhD and one held both qualifications; three apparently had no advanced degree. At the time of their involvement with Darsee, about half of the co-authors were 30–40 years old and had received their advanced degrees 5–15 years previously. Approximately equal numbers were senior and junior to this group, the latter including technicians, medical students, resident staff at teaching hospitals and junior faculty members. The senior co-authors included professors and department chairmen. (Darsee himself was 33 in 1981.)

Abundance of errors

Many of the published errors and discrepancies were presumably introduced by Darsee. We nonetheless considered all errors, major and minor, to be of interest because they might reflect on the care with which the co-authors had checked the publications that bore their names, and might reflect indirectly on the care with which they had carried out their share of the research.

All but a few of the 18 research papers[6,20-26,30-51,87-96] (Fig. 1) and many of the abstracts contain errors or discrepancies that can be recognized simply by examining them carefully. We emphasize again that many of the errors and discrepancies are minor, for example, a small discrepancy between a numerical value printed in a table and what is supposed to be the same value printed in the text. Some errors,

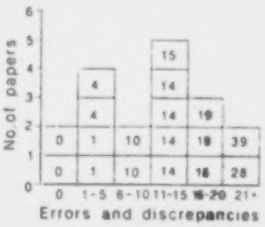


Fig. 1 Errors and discrepancies in published papers. Each box represents a single paper, and the number within the box is the sum of the errors and discrepancies for that paper.

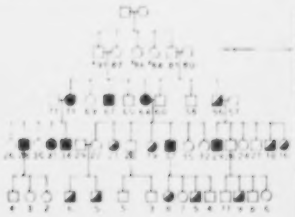


Fig. 2 Pedigree reproduced from ref. 87. The age of each family member at the start of the investigation is given below each symbol. Arrows (second row) denote those who had died before the start of the investigation. Note the 17-year-old male (row 4) with children ages 8,7,5 and 4; at the age of 8 or 9 (probably 8) he impregnated the mother of his oldest child.

however, are so glaring as to offend common sense.

We present a detailed analysis of three sample research papers from which the reader may judge the types of errors we counted. Two papers[87,88] from Emory and one from Harvard[30] are discussed below as examples of our procedure. Two of these[87,88] were eventually retracted wholly[87,88] or in part[88], but one[30] has not been retracted and is still cited as valid[87,89].

The most striking example, in ref. 87, is that in which the authors reported an arresting pedigree depicting a family with a high incidence of an unusual form of heart disease. Inspection of the pedigree (Fig. 2), reveals that a 17-year-old male had 4 children, ages 8,7,5 and 4. (It is unlikely that the father's age is a misprint: it appears in the figure and in two places in the text.) Thus the father was 8 or 9 (probably 8) when he impregnated the mother of his first child and 9 or 10 when he impregnated the mother of his second. This bizarre feature of the pedigree, which perhaps could have raised questions about the validity of the entire paper, was apparently not noticed by co-authors or referees, nor were the following unlikely groupings of ages: his sister, brother and first cousin had their first children at ages 16,15 and 15, and three women in the preceding generation had their last children at ages 41,45 and 52.

The summary of the paper gives the urinary taurine levels of the family members with congestive cardiomyopathy as ranging from 411 to 536 mg taurine per g creatinine, but the text and Table 1 of the same paper give for the same measurement 426 ± 45 mg taurine per g creatinine (mean ± standard deviation). Simple inspection suggests that these two sets of numbers cannot simultaneously be valid. Likewise, the summary of the paper states that the urinary taurine levels of the family members with mitral valve prolapse ranged from 215 to 265 mg taurine per g creatinine, whereas the text and Table 1 give 220 ± 31 mg taurine per g creatinine. These values are central to the thesis of the whole paper. The paper has three numerical discrepancies with ref. 40, and five with refs 19 and 28, abstracts which appeared before the paper and which deal with the same family. All three abstracts of this work indicated that, at the beginning of the study, all 46 members of the kindred were alive, which contradicts Fig. 1 of ref. 87, in which 3 of the 5 oldest are shown as having died before the beginning of the study.

A more subtle example of a major error is illustrated in Fig. 3, taken from a paper[88] also from Emory by six members of our sample. The figure depicts changes in haemodynamic values in five patients at various time intervals after a surgical procedure. Almost all the data in this paper[88] are contained in three figures which cannot be reconciled with each other. The mean and standard deviation for haemodynamic values in the five patients are given in our Fig. 3 (Fig. 1 in the original paper). It is easy to see that many error bars are the same size, and also that mean SVR (systemic vascular resistance) is almost constant at −160 units (the bottom of the graph) in four serial measurements. (Abbreviations are explained in the legend to Table 1.)

These results are especially implausible in the light of data given elsewhere in the paper, where it is shown that SVR and other haemodynamic values for two individual patients vary markedly in the course of time. Moreover, many standard deviations in the figure we reproduce appear on inspection too small to be reconciled with the widely scattered values shown elsewhere in the paper; this is true of SVR at 0.5 h, PVR at 1 h, SWI and SVR at 8 h, and CO at 12 h. Indeed, simple calculations show that it is virtually impossible to reconcile most of the data in the figure we reproduce with the data in the other two figures of the paper.

The introductory section in this paper[88] indicates that the study of five patients was carried out "in four uncomplicated cases of peritoneovenous shunt insertion, and in one subject who developed signs of cardiac failure and pulmonary edema", but the figure legends indicate that two of the five patients developed pulmonary oedema. Specifically, one of the four patients said to be without complications "developed clinical and radiographic signs of pulmonary edema 45 minutes after the shunt was opened" (Fig. 2 legend). The other patient, whose pulmonary oedema was more severe, is described variously as developing this complication "one hour after the shunt was opened" (Fig. 3 legend) and "within a few hours" after the shunt was opened (results section[88]).

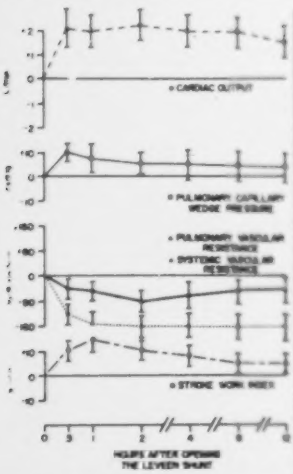
Furthermore, the statements that the changes of the mean shown in Fig. 1 after the first half hour were "very small" and that there was "very little change" appear to be contradicted by the actual maximum changes of the mean — a factor of two or more in the change from the value at time zero (corresponding to roughly a 25–60% change in the mean itself) — for three of the five parameters in the figure.

There are obvious inconsistencies between this paper[88] and a preceding abstract[90] describing the same five patients. Eight of ten values for mean or standard deviation given in the two publications are inconsistent (Table 1). This includes most of the numerical data in the abstract. Another comparison (Table 1) shows discrepancies in two out of five standard deviations between one

Table 1 Discrepancies between a paper and two abstracts

	Ref. 88	Ref. 30	Refs 20 and 88	Refs 20 and 30
Mean or s.d.	m. ± s.d.	m. ± s.d.	s.d.	s.d.
Parameter measured:				
CO	2.1 ± 0.8	2.6* ± 0.8	1.0†	0.9
PCWP	10.0 ± 3.7	12* ± 2*	4.4	1.8
PVR	−43 ± 33	−64* ± 24*	30	43‡
SVR	−123 ± 33	−192* ± 168*	199†	158
SWI	10.6 ± 4.2	11.3 ± 5.8*	4.8†	5.2‡

Ref. 88 is a full-length paper; refs 20 and 30 are abstracts. Discrepancies are indicated between columns 2 and 1 (*), columns 3 and 1 (†), and columns 4 and 2 (‡). Values in column 1 were obtained by measurements of the 0.5 h points in Fig. 1 of ref. 88. The values in column 2 were copied directly from ref. 30, except that for SVR and PVR, the R units[30] have been converted to dyn s cm^{-5} (factor of 80; ref. 144). Columns 3 and 4 contain standard deviations for 5 patients calculated from the means and standard deviations given in ref. 20 for 4 patients and the means given in refs 88 or 30 for 5 patients. The values were derived from calculations based on "population" standard deviations; calculations based on "sample" standard deviations gave similar values and conclusions. When comparing values, allowance was made for rounding of the final digit in refs 20 and 30 and for uncertainties of measurement in ref. 88. Abbreviations: CO, cardiac output; PCWP, pulmonary capillary wedge pressure; PVR, pulmonary vascular resistance; SVR, systemic vascular resistance; SWI, stroke work index. Units are given in the publications, except as noted above.



ig. 3 Reproduction of Fig. 1 from ref. 88. The raph depicts the mean and standard deviation ar changes in haemodynamic values of all five atients in the study. Note the remarkable simi- arity in size of many error bars. Note also the elative constancy of SVR (average value for ive patients) at −160 for four serial measure- nents, despite the variability of SVR (indi- vidual values for two patients) in Figs 2 and 3 of ref. 88 (not reproduced here).

abstract[29] and the other[30], and in three of five standard deviations between the abstract[29] and the paper[28]. Thus the two abstracts are inconsistent with each other and with the full-length paper.

In both research papers described above, the major errors are fundamental and call into question the validity of the paper's main conclusions. More typically, however, the errors simply reflect on the care with which the paper had been prepared.

The final example was contributed to a major journal by one of the co-authors about two months after Darsee had admitted to him and another co-author that he had forged data in another study. At the time the paper[96] was contributed, these two co-authors were carrying out, privately, a "detailed review and analysis"[114] of the extent of Darsee's forgery in their laboratory. Thus these two scientists were simultaneously publishing a paper containing the results of their collaboration with Darsee and serving in private as investigators into the extent of his forgery of data. One author stated[19] that he had "total confidence" in this particular paper, which he also said[19] had "passed special scrutiny". The Ross Committee, asked to review Darsee's work at Harvard and prepare a statement of record, affirmed the co-authors' view that this publication was accurate.

The paper[96] describes the recovery of heart muscle injured reversibly by a brief experimental occlusion of a coronary artery. One measure of heart muscle function (systolic shortening) is said in the legend to Fig. 3 not to recover fully "until after 7 days of reperfusion". This description is at least unclear, and we regard it as erroneous, since the figure itself and the data in Table 1 show full recovery at or before 7 days.

We present this as an example of an error of very minor significance, but one that we think suggests that the co-authors were not sufficiently careful in examining their own paper before publication. There are small discrepancies between the summary (p. 7152) and the text (p. 7153) with regard to ATP concentration at both 90 min and 72 h of reperfusion. There is a discrepancy between the error bar for ischaemic subendocardium at 15 min of occlusion in Fig. 1A and the corresponding number in the summary and text.

The paper also contains discrepancies with another paper[97] and with an abstract[98], both published previously, and both describing similar or identical studies on similar groups of dogs. The values of maximum negative dp/dt (all between about −4,000 and −4,300 mm Hg s^{-1}) for the dog in Fig. 3 of ref. 96 cannot be reconciled with the values (means from −2,248 to −2,562; standard deviations from 200 to 316) presumably calculated from this dog plus 6 others and listed in Table 2 of ref. 95. Also, the size of the experimental groups for measurement of ATP and creatinine phosphate at 15 min occlusion and 90 min reperfusion is different in the three publications: 14 or 21 dogs (the text is ambiguous) in ref. 95; 34 or 41 in ref. 96; and 36 in ref. 105.

Despite these differences in the numbers of dogs in the experimental groups, the means and standard errors, wherever given, are the same for all three publications. For example, in refs 95 and 96 (Figs 1 and 2, 15 min occlusion and 90 min reperfusion), there are 16 values and 16 error bars; these are identical in the two papers. Also, in the cases where numerical data are given in the text of the three publications, the numbers after rounding are the same. Such a coincidence of means and standard errors from experimental groups of different sizes is extremely implausible. The paper[96] contains additional errors and discrepancies not described here.

Among the 18 research papers, there were as many as 39 errors and discrepancies in a single research paper, with an average of about 12 per paper (Fig. 1). Of the 22 scientists who were co-authors of a research paper, 19 were co-authors of at least one research paper containing 10 or more errors or discrepancies.

The analysis of errors and discrepancies given above is based entirely on our sample of publications in scientific journals. In contrast, some of the analyses that follow depend on information given in the committee reports.

Table 2 Frequency of lapses classified by publication and by co-author

Type of lapse	No. of publications	No. of co-authors
1	28	26
2	42	21
3	7	9
4	13	6
5	7	6
6	1	4
7	8	8
8	1	1
9	1	1
10	2	3

Column 1, type of lapse as defined in Table 5. Column 2, number of publications by co-authors in which there were one or more examples of the lapse in column 1. Column 3, number of co-authors involved in each type of lapse. Some publications and some co-authors were associated with more than a single type of lapse.

Retention of data

The reports of the investigating committees showed that in almost all cases most of the co-authors had not retained the experimental or clinical data on which the publications were based. This was true of 7 of the 8 Harvard papers, which involved experiments on dogs, and of 8 of the 10 papers from Emory, 8 of which involved human subjects. [These figures appear mostly attributable to the failure to retain data by Darsee, the chief source of experimental data — Editor, *Nature*.] In the latter case, there were 7 papers for which some of the co-authors had retained neither names nor other identifying information for the human subjects. For the purposes of the present report we have disregarded all papers not involving human subjects except for the cases discussed below.

Of 14 who were co-authors of papers (all from Emory) involving human subjects, at least nine failed to retain the list of subjects (Table 2). One paper[7] is particularly interesting: a report on the incidence of mitral valve prolapse in presumably healthy young men. The paper is considered valid[17] and is heavily cited[18]. The co-authors "maintain that the data in the paper are valid although none of them has a copy of the original research results. They state that they did the study and oversaw the writing of the paper"[20]. Although Darsee was first author of the paper, the Moran report provides no information on his role in the project and, in particular, on his opportunities for altering or forging the data: possibly this in-

formation is contained in the confidential appendices to the report. Thus the only available support for the validity of this paper is based on a statement attributed by the Moran Committee to the co-authors, given without supporting data or details.

Involvement in research

In considering involvement, we confined our attention to the 18 research papers. We defined "direct involvement" to mean actual participation in the acquisition and interpretation of data, or the making of an intellectual contribution which was essential to the research. We disregarded a person's rank in the laboratory hierarchy in judging whether an author was directly involved.

Our evidence for the extent of an author's involvement came from the reports of the investigating committees. For example, according to a committee report[92], one co-author's role was to encourage Darsee and provide some of the grant support for the research. In another committee report[93] a senior scientist is described as having other duties not allowing him sufficient time for direct participation in the research laboratory and also as being seldom in the laboratory and having no direct supervisory role. In both these cases we considered that the co-authors had not met our criteria for direct involvement, in the sense of direct participation in experimental work. Supervisory functions may be exercised with a greater or lesser degree of remoteness, and we acknowledge that the evidence available is limited and that our judgements on this point are necessarily subjective.

We consider those not directly involved in the research to be "honorary" authors. Honorary authorship was a common occurrence, according to our analysis. Of 18 papers, 13 had at least one author judged by us to be honorary. There was a total of 49 authorships (that is, occurrences of an author's name on one of the 18 research papers); in 16 authorships (33% of the total), involving 6 of the 47 co-authors, the author was judged to have little or no direct involvement (Table 2). Those who were honorary authors at least once also published significantly more papers than those who were never honorary authors (3.33 ± 1.63 versus 0.71 ± 1.42, mean ± standard deviation; $P < 0.01$ by Student's two-tail t-test). Their higher rate of publication compared with co-authors who were not honorary appeared to depend entirely on their honorary papers: the difference disappeared when an author's honorary papers were not counted (0.67 ± 0.82 versus 0.71 ± 1.42). The 13 papers with at least one honorary author had more errors and discrepancies than the 5 without an honorary author (15.1 ± 10.2 versus 5.0 ± 6.6; $P = 0.067$).

Table 3 No. of dogs for which results were shared

	Ref. 53	Ref. 90	Ref. 91	Ref. 92	Ref. 94
Ref. 53	**33**	8	12	0	9
Ref. 90	8	**22**	9	0	11
Ref. 91	12	9	**76**	6	7
Ref. 92	0	0	6	**48**	0
Ref. 94	9	11	7	0	**44**

The boldface figures (on diagonal) give the total numbers of dogs used in each study. The remaining entries give the numbers of dogs for which experimental data were shared between pairs of studies. Although not shown in this table, data for the same 2 dogs were used in 4 studies[91,92,94]. Reference 134 was the source of this table, which in turn was based on information provided by one of the senior co-authors at Harvard.

Incomplete statements

In many ways the most striking of our findings was the discovery that certain research papers (7 of 18) embody statements and data in a manner that impedes the reader in forming an accurate reconstruction of the way in which experiments were actually carried out. The circumstances were these.

Five papers[91,92,94] reported studies done with dogs on the effects of experimental coronary artery occlusion on the heart and on recovery from the effects of occlusion with or without drugs. In these papers the assignment of dogs to experimental and control groups is described as formally randomized, as with the statement[94], "Twenty-five minutes after occlusion of the LAD, the remaining dogs were randomized to treated or control group on the basis of a random odd or even number generated by a computer program."

During the subsequent investigation by the NHLBI panel, however, the co-authors acknowledged that some control data had been reused[134] — in other words, that the control dogs actually included some historical controls, a use which had not been mentioned in the published papers. (Historical controls are animals that have been used as controls in a previous study[134]. The same data recorded in the earlier study are then used again as control data in the later study. The number of dogs for which data were reused is shown in Table 3.)

One co-author of three of the five papers subsequently stated[94]: "In order to economize animal usage, as well as professional and technical time, control animals which had previously been randomized were sometimes used to provide data for more than a single experiment . . . we saw no reason not to allow him [Darsee] to use his randomized controls in more than a single experiment." Another scientist, the co-author of all five papers, acknowledged the reuse of data, but stated[95] that Darsee had misled him and the other co-author about the number of dogs for which data were reused and about the number of experiments in which the data for any one dog were reused.

Reuse of data was approved by one of these co-authors in at least one paper not involving Darsee (mentioned in ref. 134), and both co-authors have subsequently defended the practice. One co-author, commenting on a study not involving Darsee that was carried out in his laboratory, acknowledged[97] that control data were reused and that the reuse was not described on publication: "Although the use [of some control data from another study] was not specified in the text [of a study not involving Darsee], the use did not bias the results of the study in any way". The other co-author defended the sharing of control results said to be randomized by stating[98] that "the practice of not specifying in the Methods Section that a control dog had been used in more than one study, while certainly not ideal, is common," and also that "the practice of utilizing controls in more than a single experiment is not uncommon and, *although it should be*, it is not always spelled out in published papers" (emphasis added).

In a series of experiments in which the objective is to study the effects of various interventions on animals subjected to a certain procedure, in this case arterial occlusion, the ideal is that described in this group of five papers: all animals would be subjected to arterial occlusion, and afterwards assigned randomly either to the control group or to one or other of the experimental groups.

Although the use of historical controls may have several practical advantages, not least those of economizing in the use of animals, this practice is clearly less satisfactory than the ideal; for example, the experimenter's foreknowledge could systematically affect the outcome of experiments. The fact that Darsee may not have selected historical controls at random, as required by the laboratory's policy described to the NHLBI panel, is irrelevant to our criticism, which is that these papers leave the reader with the impression that control animals were selected by a randomization procedure that most readers, we judge, would consider superior to that actually used.

We emphasize that we do not criticize the use of historical controls, which may or may not have been appropriate, but merely the misleading published description. Thus the use of historical controls is reconcilable with the statement in the

paper quoted above that "after occlusion . . . the remaining dogs" were randomly assigned to different groups only if "remaining dogs" is understood to refer to a number of historical controls.

It would appear from the subsequent statements of the two co-authors that they knew of and authorized the use of historical controls in the laboratory, as well as the failure to describe that use on publication. The described randomization procedure in any study was supposed to apply to all dogs in that study, but in fact did not apply to the historical controls when their data were reused. We conclude that the papers contained an incomplete and misleading description of experimental procedures.

Furthermore, the experiments reported in the five papers included procedures, such as arterial catheterization and administration of certain drugs or dyes, that differ from one study to another (Table 4). As a result, most of the historical controls had been subjected to procedures in an earlier study that were different from those described for the later studies. Thus the published descriptions of these later procedures, which are described explicitly and in detail as applying to all dogs, do not apply in some respects to the historical controls.

One author, writing to the NIH about the sharing of results, stated[108] that "the experimental conditions were similar" for the studies in question. The other referred[107] to studies (not involving Darsee) in which "experimental conditions are the same"; he has also indicated in correspondence with us that the procedures described in the five papers were similar enough in some cases so that the use of historical controls was justified. Differences in experimental conditions are summarized in Table 4.

Some might consider that in most cases shown in Table 4, the data from dogs in one study would be unsuitable as reused data for another study, but the facts enabling readers to form a judgment on this point were not given in any of the five papers. We believe that the descriptions of experimental procedures did not apply to the reused data obtained in previous studies and that these descriptions were therefore misleading.

A final point may be related to reuse of data: one of the five papers states[91] that "all dogs [were subjected to an experimental procedure, which is described]. Ventricular fibrillation or standstill was then induced by a barbiturate overdose, and the hearts were excised rapidly and placed on ice." Yet the co-authors have said (cited in ref. 134) that some of the same dogs were subsequently used for the 4-hour and 6-hour occlusions in other experiments. As the NHLBI panel report politely observes[109], "This would appear to be impossible". We have assumed that the two co-authors in our sample were unaware of this anomaly at the time of publication.

The misleading statements noted above appeared in five papers co-authored by two members of our sample. In addition to these five papers, there were two others (refs 95 and 96; see next section) containing statements that were imprecise or misleading. In total, there were 18 instances of statements made by 6 co-authors whose effect would have been, in our opinion, to mislead the reader.

Unacknowledged republication

We were surprised to discover several instances in which the same data were republished without adequate reference to that circumstance. The most striking example was a paper[95] containing data later published almost in their entirety as a major part of a subsequent paper[96] which appeared in a different journal. Although the second paper referred to the first by means of an index number in the text, there is no explicit statement in the second paper that the new data are a continuation of an experiment already published.

The subject matter of both papers concerns the recovery of heart muscle injured by temporary surgical occlusion of a coronary artery in dogs, described by the authors as a model with "important clinical implications for patients" with certain kinds of heart disease. The first paper states that "Seven dogs were also subjected to 15 min of LAD occlusion followed by 3 days of reperfusion", which is not false but which does not reveal that the same 7 dogs were actually reperfused for 14 days. The references in the second paper to the first would not suggest to those who were not familiar with the first that the research about to be described was a continuation of an experiment already published. In the second paper's opening paragraph, which is reproduced overleaf, the first paper is cited twice, but neither citation of the first paper provides a hint that some of the same data are being republished and that some of the same animals were used.

On neither paper was Darsee the first or last author, and it seems that the co-authors could have known of the statements that were imprecisely misleading. For the journal to which ref. 96 was submitted (*Proceedings of the National Academy of Sciences*), a contributing author customarily by-passes the refereeing system, which might have discovered the unacknowledged republication of data; instead the contributing author assumes "responsibility for the propriety and scientific standards"[110] of the manuscript he submits.

Table 4 Differences between experimental procedures in five papers

	Ref. 53	Ref. 90	Ref. 92	Ref. 94	Ref. 91
(1) Coronary artery occlusion (time)	2 - 6 h	6 h	1 or 3 h	4 - 6 h	1 min - 2h
(2) Randomization before/after coronary artery occlusion	After	After	Before	After	Before
(3) Time from removal of coronary artery clamp to end	~1 min - 4h	Not removed	5 or 72 h	Not removed or 2h	6 h
(4) Antiarrhythmic agent:	Lido	Lido	Lido	Proc'de	None
(i) Initial bolus inj., mg per kg	1.5	1.5	1.5	3	None
(ii) Further intermittent inj.	Yes	No(?)	No(?)	No(?)	No
(iii) Continuous infusion	No	No	No	Yes	No
(5) Dye administered initially	Fln	MeB,GV	Fln,GV	Fln	Fln,GV
(6) Dye injected terminally	TS,MeB	MeB	MeB	MonB	MonB
(7) Drug or NaCl soln, ml per kg	0(?)	1	1	2	0(?)
(8) Catheter in aorta/left carotid	Carotid	Carotid	Carotid	Aorta	Carotid
(9) A set of multiple procedures, (a) - (e), listed below	No	No	Yes	No	Yes

There are substantial differences in experimental procedures for at least nine of the ten pairs of papers. The procedures in item 9 include: (a) insertion of micromanometer through stab wound in ventricular apex, (b) application of coronary cuff, (c) repeated brief reocclusion of coronary artery, (d) implantation of crystal transducers in myocardium, and (e) delay in examination of heart after death (for flowmeter calibration). Abbreviations: Lido, lidocaine; Proc'de, procainamide; Fln, fluorescein (injection); MeB, methylene blue (injection); GV, gentian violet (topical); TS, thioflavine S; MonB, monastral blue. Information in the table comes from the papers themselves[53,90,91,92,94]; see also ref. 134, appendix 5.

Opening paragraph of ref. 96:

Previous studies from this laboratory have shown that brief (15 min) coronary occlusions followed by reperfusion do not cause necrosis but do result in prolonged abnormalities in myocardial biochemistry, function, and ultrastructure (1 – 3). At 72 h after release of a 15-min coronary occlusion, the ATP concentration in reperfused previously ischemic myocardium was significantly reduced (by 22% of normal), the percentage of systolic shortening of reperfused left ventricular segments was reduced (by 42% of normal), and ultrastructural abnormalities were present, despite the absence of necrosis (1, 3). However, it was not known whether these changes ever were reversed to normal after longer periods of reperfusion and, if so, what the time course of the recovery is. Recently there has been renewed interest in the treatment of acute myocardial infarction by means of coronary reperfusion, either by injecting fibrinolytic agents directly into the obstructed coronary artery or by coronary revascularization (4, 5).

This is the opening paragraph, printed here verbatim, of a research paper[96], discussed in the text. The two occurrences of reference 3 in the paragraph above are the only references in this paper[96] to the authors' own previous study[97] (see text).

A common example of republication involves abstracts of research submitted to scientific societies. Of the 88 abstracts, only 47 were published only once; all the rest involve duplicate or triplicate publication. This practice is apparently followed in many fields of biomedical science, where an abstract may be prepared for one meeting of a society and then submitted, perhaps in an amended form, to a second meeting. We nevertheless suspect that the extent of this practice may not be widely appreciated. We discuss four cases below but have not included abstract data in our tables.

We note that there were four pairs of Darsee abstracts in which the texts were nearly identical, but whose titles had apparently been changed. For example, an abstract[71] entitled "Can tetrazolium stains reliably identify myocardial infarction prior to definite histologic necrosis?" was also presented[72] under the title "Early pathologic detection of acute myocardial infarction". Sometimes the alteration consists of the substitution of synonyms[70,72]: "Persistent myocardial abnormalities following brief periods of temporary coronary occlusion not associated with necrosis" also appears as "Prolonged metabolic, functional, and ultrastructural abnormalities following transient myocardial ischemia without infarction".

In each pair the authors are the same, and the text and numerical values are identical through most of the abstract. The publication dates were generally within about a month of one another, as were the scheduled presentation dates. Although it could be argued that it may be necessary to tailor the titles of certain abstracts to make them acceptable for oral presentation before the different societies, we consider that the change in title has the effect of making the pair of abstracts, if listed only by title, appear different to the reader, whereas in fact the contents are almost the same.

There were 5 instances of unacknowledged republication of data. They appeared in publications, two from Emory and seven from Harvard, having a total of 8 co-authors (Table 2).

Lapses from standards

We were surprised to find that there is apparently no generally accepted written code of conduct for scientific research[100-129]. We think that there is nonetheless an unwritten code of which scientists are aware and to which they profess adherence. We have based the analysis that follows on standards that we believe are generally accepted — accepted in the sense that they are generally understood, not necessarily in the sense that they are always complied with.

In Table 5 we attempt to place in ten categories the departures from generally accepted standards that we observed in our sample; other categories would probably be more appropriate for other studies. We found it useful to divide our categories into two groups, A and B, with those of group A explicable simply by carelessness or excessive haste, while those of group B appeared to be more serious in some sense. We do not mean to suggest that certain departures from accepted practice which appear in group B may not be due solely to carelessness and haste; undoubtedly this is so in some cases. Furthermore, we recognize that the distinction between type A and type B is necessarily subjective and that other observers might classify the lapses we have found differently.

We emphasize that the instances on which Fig. 4 is based were supported by evidence we considered compelling, but not necessarily conclusive in each case. We acknowledge that our categories must be subjective.

Frequency of lapses

Of the 47 co-authors in our sample, 31 (66%) were authors of publications containing lapses of type A — a result that is the same whether or not honorary authorships are included. Furthermore, 13 (28%) were involved in type B lapses. Only 12 of the 47 were found not to have been involved in lapses of either type. Most of the 12 were authors of abstracts only; just 2 of the 12 were co-authors of a research paper.

Some scientists apparently do not regard abstracts as publications, even though they are archived and cited and are sometimes used to establish priority[100]. If the 88 abstracts and 3 chapters are excluded from consideration, the frequency of type A and B lapses is about the same or higher. Of the 22 scientists who were co-authors of the 18 research papers, 21 (95%) were involved with papers containing lapses of type A, and 6 (27%) were involved in lapses of type B.

Discussion

It is obvious that the lapses from generally accepted standards of research described here have at least the potential for interfering with the accuracy of the scientific literature or for harming the scientific enterprise in some other way. We also think that the instances described above represent departures from the generally accepted standards of scientific research. We recognize that the decision as to what constitutes a departure from generally accepted standards is necessarily subjective. Presumably some scientists have views different from those expressed here. We regard it as a matter of the greatest importance that these differences be discussed publicly.

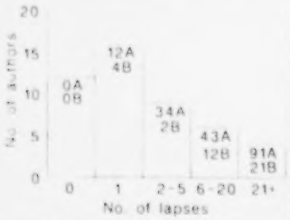
Table 5 Lapses from generally accepted standards

Type A

(1) Presence of errors (usually in numerical values).
(2) Inconsistency with a research group's previously published data.
(3) Failure to retain names or identifying numbers for human subjects included in a published paper.
(4) Honorary authorship (see text).

Type B

(5) Statements that are misleading.
(6) Misleading citations of a previous paper containing data being republished, so that the true relationship between the two papers is obscured.
(7) Publication of very similar abstracts under very different titles.
(8) Failure to utilise unique knowledge that makes possible the recognition of certain serious errors in published work — errors not discoverable by anyone else.
(9) Failure to acknowledge the source of a substantial amount of research data received from someone else.
(10) Failure to take appropriate action after receipt of a complaint, later shown to be well founded, that a colleague may be involved in questionable data collection.



ig. 4 Authors and their lapses. The histogram ives the number of authors in our sample have ne lapses within the specified range. The numers give the total of type A and type B lapses r all authors in that range.

The abundance of errors that we bserved in our sample obviously interfers with the accuracy of the scientific literature. Every scientist has, at a minimum, n obligation to ensure that what is pubished under his name is accurate.

There is a notable absence of explicit ublished standards for the retention of ata. We believe that it would be generals accepted that data should remain accesable for as long as questions and criticisms re likely to be raised about what has been ublished. For most of the papers in our ample, data were not available. The data we give in Table 4 use only those cases where data involving human subjects had ot been retained. We suggest that a reasonable minimum requirement is that each author of a paper involving human subjects should retain at least a list of the subjects' names or the hospital's identification numbers of patients. This information would permit access to hospital records and thus in many cases allow reconstruction of data in the paper.

The reader may ask: What harm is done by honorary authorship? Indeed some of our colleagues have argued that the custom of routinely placing the name of a senior scientist, usually the head of the laboratory, on a paper — regardless of his contribution — is widely followed and does no harm. We disagree, as have others[156,157,158]: honorary authorships falsify the assignment of responsibility for published research and increase the likelihood that inaccurate data will be published. The honorary author is in a poor position to judge the validity of the work, yet he often lends a prestige that may lull other co-authors, the reviewers or the readers into uncritical and inappropriate acceptance.

Similarly, the harm done by unacknowledged republication of data may not be immediately obvious. But it disguises the true relationship between published papers and yields no benefits other than for the authors.

In some ways the most disturbing of our findings is the discovery that certain of the publications (7 out of 18) embody statements and data in a manner that impedes the reader in forming an accurate reconstruction of the way in which experiments were carried out.

Representativeness

The evidence presented here indicates a surprisingly high frequency of lapses from accepted standards. One possible explanation is the association of the co-authors of the publications in our sample with Darsee: he may have induced them to behave in ways not typical of them. A second possibility is that the co-authors are not representative of their field. A third possibility is that lapses from accepted standards may be unusually common in the specialty in which these scientists worked, in clinical research generally or in research carried out at medical schools or by physicians. (The co-authors of the publications in our sample were mainly physicians, and the research was carried out exclusively in medical schools.) A fourth and disturbing possibility is that lapses from accepted standards may be more common among biomedical scientists than is currently appreciated.

Other studies will be needed to determine whether our results are representative of other groups of scientists.

Causes

Extraordinary emphasis is commonly placed on the sheer quantity of publications[159,160,170]. The director of one of the world's leading research institutions (but not an institution with which Darsee was ever affiliated) sent a signed, official memorandum to junior colleagues proposing incentives for publication: "There is no demand that these be literary masterpieces in first line journals; journeyman works for publication in second, third or fourth line archival publications will be quite satisfactory. Upon proper completion and submission of [two] manuscripts, [a technician]'s appointment will be extended to April 1, 1984. During that time it is expected that an additional manuscript on [subject] will be completed and submitted. If so, the period of employment will be extended an additional three months and again an additional manuscript on [subject] is an anticipated result of the extended employment." On another occasion a director of research said in all seriousness, "No research that results in a single paper is worthwhile."

Those promoting Darsee or recommending his promotions were well aware that his publications were unusually numerous[146]. But apparently none of these supporters, who were also his co-authors, made the effort that would have shown that many of the papers had serious flaws. How often does this occur?

Other causes of lapses from accepted standards may be the extreme competition for research grants, large research teams in which the leader is not directly involved in the research and the absence of an accepted code of conduct for scientific research.

In this paper, we have deliberately avoided a discussion of the cases of acknowledged fraud which have come to light in recent years, concentrating instead on defects of the scientific literature generally regarded in the scientific community as regrettable but less serious. Outright fraud, which might conveniently be called a lapse of type C in our nomenclature, is presumably rare. May not the scientific community be over-sanguine in attaching less importance to lapses such as we have found, which are generally unremarked but which debase the scientific literature?

Testing the system

The results of our study naturally raise the question: what fraction of papers in the biomedical sciences are not supported by primary data at the time of publication? The question cannot be answered with the papers in our sample, since most are known to be partial or complete forgeries and thus are atypical.

We suggest a study to answer the question: close examination of a random sample of published papers. Probability of selection could be uniform or could be weighted by a variable such as number of authors or cost of research. Papers would be examined by either of two methods: careful examination of the paper itself for errors and discrepancies by appropriate experts following an established procedure ("external audit") or examination of the primary data on which the paper is based, which would of course require access to laboratory records ("internal audit"). The advantages of the external audit are its non-invasive character and presumably lower cost; the advantage of the internal audit is its completeness.

What would this cost? This depends on the precision required and on the actual fraction of defective papers. We assume for the sake of illustration that an external audit could be done for about 3% of the cost of the research on which a paper is based, and an internal audit for about 10%. If the actual fraction of papers that could not be substantiated were as high as 15% and if the answer were required with a standard error of 7%, about 25 papers would have to be examined, which in the case of internal audits would altogether cost only about 2.5 times as much as one average paper. If, on the other hand, auditing 100 papers failed to disclose a single paper that could not be substantiated, this would show with $P = 0.05$ that the actual fraction of papers unsupported by primary data at the time of publication is less than 3%.

In addition to the financial costs of examining the practices of scientists, there are other costs which may be more serious

in the long run. Examination of scientific practices could cause unwarranted harm to individual scientists. Systematic examination of scientific practices might even weaken the fabric of trust that is essential to the functioning of science. Of all human endeavours, science is one of the most successful — prodigious in benefits, low in cost. But science, vulnerable to abuse from within by its practitioners, is perhaps even more vulnerable to harm by regulation, and at some point the cost of further regulation will outweigh the benefits.

Scientists have to an unusual degree been entrusted with the regulation of their own professional activities. Self-regulation is a privilege that must be exercised vigorously and wisely, or it may be lost. □

Walter Stewart and Ned Feder are in the National Institutes of Health, Bethesda, Maryland 20892, USA.

1. *Notre Dame Sci. Q.* 7, no. 2, 14-17 (1969).
2. *Notre Dame Sci. Q.* 8, no. 1, 9-11 (1969).
3. *Circulation* **56**, III-218 (1977).
4. *South. med. J.* **70**, 1249 (1977).
5. *Arch. intern. Med.* **137**, 1639 (1977).
6. *Ann. intern. Med.* **89**, 867-870 (1978).
7. *Clin. Res.* **26**, 3A (1978).
8. *Clin. Res.* **26**, 226A (1978).
9. *Clin. Res.* **26**, 226A (1978).
10. *Chest* **74**, 334-335 (1978).
11. *Chest* **74**, 334 (1978).
12. *Circulation* **58**, II-237 (1978).
13. *Circulation* **58**, II-235 (1978).
14. *Circulation* **58**, II-229 (1978).
15. *Circulation* **58**, II-174 (1978).
16. *Circulation* **58**, II-42 (1978).
17. *Circulation* **58**, II-23 (1978).
18. *Clin. Res.* **26**, 771A (1978).
19. *Clin. Res.* **26**, 762A (1978).
20. *Clin. Res.* **26**, 759A (1978).
21. *Clin. Res.* **26**, 745A (1978).
22. *New Engl. J. Med.* **298**, 221-222 (1978).
23. *Am. J. Med.* **64**, 1069 (1978).
24. *Circulation* **59**, 492-497 (1979).
25. *Circulation* **59**, 619-622 (1979).
26. *New Engl. J. Med.* **300**, 877-882 (1979).
27. *Am. J. Cardiol.* **43**, 367 (1979).
28. *Clin. Res.* **27**, 274A (1979).
29. *Clin. Res.* **27**, 265A (1979).
30. *Clin. Res.* **27**, 264A (1979).
31. *Clin. Res.* **27**, 230A (1979).
32. *Clin. Res.* **27**, 161A (1979).
33. Abstract of report presented in annual Lloyd Hyde Medical House Staff and Fellows Research Program, Emory University School of Medicine (1979).
34. *Clin. Res.* **27**, 573A (1979).
35. *Clin. Res.* **27**, 561A (1979).
36. *Clin. Res.* **27**, 561A (1979).
37. *Chest* **76**, 371 (1979).
38. *Chest* **76**, 350 (1979).
39. *PAC E 2*, A43 (1979).
40. *Circulation* **60**, II-157 (1979).
41. *Circulation* **60**, II-149 (1979).
42. *Circulation* **60**, II-124 (1979).
43. *Circulation* **60**, II-90 (1979).
44. *Circulation* **60**, II-41 (1979).
45. *Circulation* **60**, II-38 (1979).
46. *Clin. Res.* **27**, 726A (1979).
47. *New Engl. J. Med.* **301**, 443 (1979).
48. *South. med. J.* **72**, 174-180 (1979).
49. *Circulation* **60**, II-292 (1979).
50. *Ann. intern. Med.* **92**, 735-741 (1980).
51. *Am. J. Cardiol.* **46**, 607-612 (1980).
52. *Am. J. Cardiol.* **46**, 613-618 (1980).
53. *Am. J. Cardiol.* **46**, 800-806 (1980).
54. *Am. J. Cardiol.* **45**, 892 (1980).
55. *Am. J. Cardiol.* **45**, 476 (1980).
56. *Am. J. Cardiol.* **45**, 395 (1980).
57. *Clin. Res.* **28**, 257A (1980).
58. *Clin. Res.* **28**, 163A (1980).
59. *Clin. Res.* **28**, 163A (1980).
60. *Clin. Res.* **28**, 653A (1980).
61. *Chest* **78**, 513 (1980).
62. *Chest* **78**, 513 (1980).
63. *Circulation* **62**, III-318 (1980).
64. *Circulation* **62**, III-317 (1980).
65. *Circulation* **62**, III-177 (1980).
66. *Am. J. Cardiol.* **45**, 394 (1980).
67. *Fedn Proc.* **39**, 1112 (1980).
68. *Fedn Proc.* **39**, 986 (1980).
69. *Clin. Res.* **28**, 536A (1980).
70. *Clin. Res.* **28**, 163A (1980).
71. *Clin. Res.* **28**, 174A (1980).
72. *Clin. Res.* **28**, 615A (1980).
73. *Clin. Res.* **28**, 614A (1980).
74. *Clin. Res.* **28**, 610A (1980).
75. *Clin. Res.* **28**, 609A (1980).
76. *Circulation* **62**, III-314 (1980).
77. *Circulation* **62**, III-246 (1980).
78. *Circulation* **62**, III-230 (1980).
79. *Circulation* **62**, III-196 (1980).
80. *Circulation* **62**, III-80 (1980).
81. *Circulation* **62**, III-76 (1980).
82. In *Clinical Methods: The History, Physical and Laboratory Examinations* (eds Walker, H. K., Hall, W. D. & Hurst, J. W.) 1122-1161 (Butterworth, Boston, 1980).
83. In *The Heart, Update III* (ed. Hurst, J. W.) 1-8 (McGraw-Hill, New York, 1980).
84. In *Heart Disease: A Textbook of Cardiovascular Medicine*, Vol. 2 (ed. Braunwald, E.) 1517-1582 (Saunders, Philadelphia, 1980).
85. *Primary Cardiol.* **6**, 101-109 (1980).
86. *Primary Cardiol.* **6**, 112-118, 122 (1980).
87. *New Engl. J. Med.* **304**, 129-135 (1981).
88. *Ann. Surg.* **194**, 188-192 (1981).
89. *Circulation Res.* **49**, 1017-1028 (1981).
90. *Circulation* **63**, 29-35 (1981).
91. *Am. J. Physiol.* **240**, H399-H407 (1981).
92. *J. clin. Invest.* **68**, 225-239 (1981).
93. *Archs Path. Lab. Med.* **105**, 403-406 (1981).
94. *Am. J. Cardiol.* **48**, 702-710 (1981).
95. *Am. J. Physiol.* **241**, H591-H599 (1981).
96. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7152-7156 (1981).
97. *Clin. Res.* **29**, 444A (1981).
98. *Clin. Res.* **29**, 428A (1981).
99. *Chest* **80**, 398 (1981).
100. *Circulation* **64**, IV-296 (1981).
101. *Circulation* **64**, IV-210 (1981).
102. *Circulation* **64**, IV-69 (1981).
103. *Circulation* **64**, IV-68 (1981).
104. *Fedn Proc.* **40**, 485 (1981).
105. *Clin. Res.* **29**, 562A (1981).
106. *Clin. Res.* **29**, 270A (1981).
107. *Clin. Res.* **29**, 270A (1981).
108. *Clin. Res.* **29**, 184A (1981).
109. *Clin. Res.* **29**, 184A (1981).
110. *Clin. Res.* **29**, 184A (1981).
111. *Z. Kardiologie* **70**, 317 (1981).
112. *J. nucl. Med.* **22**, P53 (1981).
113. *Chest* **80**, 365-366 (1981).
114. *Circulation* **64**, IV-200 (1981).
115. *Circulation* **64**, IV-236 (1981).
116. *Circulation* **64**, IV-222 (1981).
117. *Circulation* **64**, IV-178 (1981).
118. *Circulation* **64**, IV-122 (1981).
119. *Circulation* **64**, IV-99 (1981).
120. In *The Heart, Update IV* (ed. Hurst, J. W.) 219-286 (McGraw-Hill, New York, 1981).
121. *Am. Heart J.* **102**, 807-808 (1981).
122. *Am. Heart J.* **101**, 124-126 (1981).
123. *J. cardiovasc. Med.* **6**, 705-715 (1981).
124. *Acta med. scand., Suppl.* **651**, 123-132 (1981).
125. *J. cardiovasc. Med.* **6**, 1173-1177, 1181 (1981).
126. *Am. J. Cardiol.* **49**, 1023 (1982).
127. *New Engl. J. Med.* **308**, 1419 (1983) (retraction).
128. *Ann. intern. Med.* **99**, 275-276 (1983) (retraction).
129. *Circulation Res.* **53**, 837 (1983) (retraction).
130. Last, J. M. (ed.) *A Dictionary of Epidemiology*, 23, 94 (Oxford University Press, New York, 1983).
131. Ross, R. S. *et al.* Report of an ad hoc advisory committee to the dean of the Harvard Medical School on dishonesty in scientific research. *Harvard University Gazette* **77**, 1, 11-12 (29 January 1982).
132. Hall, E. C. *et al.* Report of ad hoc committee to evaluate research of Dr. John R. Darsee at Emory University. [the 'Moran report'], 1 - 32 (Emory University, 28 March 1983); cf. *Minerva* **23**, 276 - 284 (1985).
133. Moran, N. C. Addendum to report on research of Dr. John R. Darsee at Emory University School of Medicine, 1 - 3 (Emory University, 5 May 1983).
134. Morgan, H. E., Klocke, F. J., Shepherd, J. T. & Wildenthal, C. K., Report of the National Heart, Lung, and Blood Institute special panel to review alleged misconduct at Brigham and Women's Hospital/Harvard Medical School, Boston, Massachusetts, 1 - 36 and 7 Appendices (29 June 1982).
135. Knox, R. *J. Am. med. Ass.* **249**, 1797-1807 (1983).
136. Knox, R. A. *J. Am. med. Ass.* **249**, 2867-2876 (1983).
137. Knox, R. A. *Boston Globe*, 41, 46 (23 May 1983).
138. Broad, W. J. *Science* **215**, 476-482 (1982).
139. Broad, W. J. *Science* **215**, 874-876 (1982).
140. Culliton, B. J. *Science* **220**, 31-35 (1983).
141. Culliton, B. J. *Science* **220**, 936 (1983).
142. Heymsfield, S. B. & Glenn, J. F. *New Engl. J. Med.* **308**, 1403 (1983).
143. [Braunwald, E.] *Proc. natn. Acad. Sci. U.S.A.* **79**, 6980 (1982).
144. Milnor, W. R. *Hemodynamics*, 17 (Williams and Wilkins, Baltimore, 1982).
145. *Science Citation Index, Source Index* (Institute for Scientific Information, Philadelphia, 1979-86).
146. Braunwald, E., "Comments concerning NHLBI panel's report and range of possible NIH actions", Memorandum to NIH staff members, 1-36 and 6 Appendices (21 September 1982).
147. Kirner, R. A., Memorandum to W. F. Raub, NIH, 1-21 (14 September 1982).
148. "Information To Contributors", *Proc. natn. Acad. Sci. U.S.A.* **78**, ix (January 1981).
149. Krevans, J. R. *et al. J. med. Educ.* **57**, 895-902 (1982).
150. Pigman, W. & Carmichael, E. B. *Science* **111**, 643-647 (1950).
151. Clapp, J. *Professional Ethics and Insignia*, 1-851 (Scarecrow Press, Metuchen, New Jersey, 1974).
152. "Yale Policy Statement on Collaborative Research", *Yale Weekly Bull. Calendar* **11**, 9 (13-20 September 1982).
153. Hastings Center Staff (eds) *The Hastings Center's Bibliography of Ethics, Biomedicine, and Professional Responsibility*, 10-11 (University Publications of America, Frederick, Maryland, 1984).
154. Danforth, W. H. *et al. Report of the Association of American Universities Committee on the Integrity of Research*, 1-6 (Association of American Universities, Washington, DC, [April 1983].
155. *The Chemist's Creed*, American Chemical Society (ACS, Washington, DC, 14 September 1965).
156. *Honor in Science*, 1-38 (Sigma Xi, New Haven, Connecticut, 1984).
157. Chalk, R., Frankel, M. S. & Chafer, S. B. *Professional Ethics in the Scientific and Engineering Societies*, 1- 224 (American Association for the Advancement of Science, Washington, DC, 1980).
158. Huth, E. J. *Ann. intern. Med.* **104**, 266-267 (1986).
159. Huth, E. J. *Ann. intern. Med.* **104**, 266-274 (1986).
160. Stossel, T. P. *New Engl. J. Med.* **313**, 123-126 (1985).
161. Cannon, W. B. *The Way of an Investigator: a Scientist's Experience in Medical Research*, 93 (Norton, New York, 1945).
162. [Huth, E. J.] *Ann. intern. Med.* **97**, 613-614 (1982).
163. Relman, A. S. *New Engl. J. Med.* **308**, 1415-1417 (1983).
164. Croll, R. P. *Perspect. Biol. Med.* **27**, 401-407 (1984).
165. Majerus, P. W. *J. clin. Invest.* **70**, 213-217 (1982).
166. Huth, E. J. *Ann. intern. Med.* **104**, 257-259 (1986).
167. Broad, W. & Wade, N. *Betrayers of the Truth* (Simon & Schuster, New York, 1982).
168. [Huth, E. J.] *Ann. intern. Med.* **99**, 266-267 (1983).
169. Burman, K. D. *Ann. intern. Med.* **97**, 602-605 (1982).
170. Gjerde, C. L. & Colombo, S. E. *J. med. Educ.* **57**, 157-162 (1982).
171. Angell, M. *Science* **219**, 1417-1418 (1983).
172. Mahoney, M. J. *Scientist as Subject: the Psychological Imperative* 79-82 (Ballinger, Cambridge, Massachusetts, 1976).
173. Jackson, C. I. & Prados, J. W. *Am. Scient.* **71**, 462-464 (1983).
174. Gratzer, W. *Nature* **302**, 774-775 (1983).
175. Stetten, D., Jr *Science* **226**, 1374, 1376 (1984).
176. Peterdorf, R. G. *Ann. intern. Med.* **104**, 252-254 (1986).
177. Wooff, P. K. *Ann. intern. Med.* **104**, 254-256 (1986).
178. Butler, J. C., III. *Ann. intern. Med.* **104**, 259-261 (1986).
179. Angell, M. *Ann. intern. Med.* **104**, 261-262 (1986).

● **Some editorial changes have been made in this manuscript without the consent of the authors. A reply from Braunwald follows on page 215.**

NATURE VOL. 325 15 JANUARY 1987

On analysing scientific fraud

from Dr Eugene Braunwald, Harvard Medical School

ANYONE who has been a victim of scientific fraud or who has had close contact with such an episode must be sobered by the event and must ask how the fraud could have been prevented or detected earlier. The costs are incalculable, not only for the persons and institutions involved, but for science as a whole. An objective analysis of research practices which offered some insight into the nature of scientific fraud and the manner in which it could be prevented or detected would be an important contribution to the public discussion of this critical issue.

I agree that many of the issues Stewart and Feder raise are important and deserve attention and thoughtful analysis. When specific incidents of scientific misconduct are discussed, however, it is essential that the facts be dealt with accurately and objectively. Unfortunately, Stewart and Feder do not provide such a discussion of the events surrounding the Darsee case insofar as they occurred at the Cardiac Research Laboratory at Harvard and the Brigham and Women's Hospital.

As the paper has been widely circulated in many different versions and its contents have been released to the lay press before publication, I am pleased to have the opportunity to reply to the paper with respect to the work done by the 23 coauthors at Harvard.

John Darseee committed serious fraud in at least three (not two) universities over a period of at least 12 years. The premise implicit in Stewart and Feder's paper is that it is possible to separate clearly *this* fraud committed by Darsee from the professional behaviour of his coauthors. In fact, the paper repeatedly and unfairly connects Darsee's fraud at Harvard to his coauthors there through a process of guilt by association. As a result, many of Stewart and Feder's allegations are misattributions of Darsee's wrong doing to his coauthors. Other allegations involving the papers from Harvard — all of which were retracted in 1982 — are insubstantial or simply wrong and have nothing to do with scientific misconduct.

Allegations of errors

Stewart and Feder assert that there were errors in the Harvard papers although the specifics and location of those purported errors in the papers have not been provided despite numerous requests. This assertion is highly dubious as the few alleged errors which Stewart and Feder actually identify prove, upon careful scrutiny, not to be errors at all or to be wholly trivial errors which most scientists would agree do not rise to the level of a "lapse" in research standards.

Stewart and Feder discuss one Harvard paper, Paper 96, in which they purport to have identified five errors. The first involves a supposed semantic inconsistency concerning when systolic shortening of the previously ischaemic heart returned to normal. The legend to a figure in Paper 96 stated that recovery of systolic shortening did not occur until *after* 7 days, but Stewart and Feder claim that the figure and a table reflected recovery *at* 7 days. Paper 96 makes clear that the data presented in the figure and the table were obtained at the end of 7 days and the phrase "after 7 days" should not mislead anyone. If this represents an error in phrasing at all, I agree with Stewart and Feder that it is of "very minor significance".

Equally insignificant are two claimed discrepancies between values reported in the summary and text of Paper 96, which Stewart and Feder concede are "small". Whether these small differences resulted from the different methods by which these figures could be calculated (on an experiment-by-experiment basis versus a group basis) or from typographical errors, they did not affect the findings of Paper 96 and would not mislead any reader. The general understanding of scientific fraud is hardly advanced by a discussion which hinges upon typographical errors and similar quibblings.

Stewart and Feder also point to a supposed discrepancy between the value of negative dp/dt for one dog in Paper 96 and for six dogs in Table 2 of Paper 95, which had been submitted for publication six months earlier. These alleged errors can be identified only with hindsight, for there is nothing inherently suspicuous about the error bars or the dp/dt values. Although these may ultimately be inaccurate as a result of Darsee's fraud, it was not a questionable practice for the Harvard coauthors *at the time* to have accepted Darsee's calculations. Any other course would require coauthors of each paper to recalculate individually every value reported, denying researchers the considerable benefits of a reasonable division of labour. Indeed, in the type of experiment reported, although the absolute values of the cardiovascular variables may differ considerably among different animals, they tend to change in a similar manner over time, and it should not be surprising that the error bars are similar.

Fourthly, the purported discrepancy between the number of experiments in the two papers and the similarity of the results relates to the fact that the number of animals studied at various times differed. The results in those animals reperfused for the shortest period were similar in the two papers. Smaller numbers were reperfused for longer periods. Stewart and Feder's suggestion that those values and error bars could not be correct is unfounded. At most, there was ambiguity about which dogs provided biochemical data at different times. This is not a "lapse of standard".

Lastly, Stewart and Feder claim an inconsistency between an error bar in a figure and the numerical values presented in the text and summary. That and other figures in Paper 96 reflect numbers in the text within the technical limitations of artwork in scientific papers.

Retention of data

I certainly agree with Stewart and Feder about the need to retain research data, but they convey a totally inaccurate picture about this practice in the Harvard laboratory because they fail to mention that the NHLBI panel expressly found that only Darsee failed to retain primary data, and that the other Harvard researchers did retain original records. The NHLBI report stated, "With the exception of Dr Darsee, recent trainees [in the Harvard laboratory] have retained original records in customary and acceptable fashion". The laboratory policy on retention of data was clear; Darsee's failure to retain data was in direct contravention of the policy. Stewart and Feder would no doubt agree that it is proper practice in describing results to provide all of the relevant information. They did not have direct information about retention of data; the NHLBI Panel did. It is unfair to criticize the Harvard coauthors about this issue while withholding from the reader relevant information obtained by the panel.

Involvement in research

I fully agree with Stewart and Feder's general conclusion about the corrosive influence of "honorary coauthorship", but, as they admit, their assessment of this issue is based on subjective inferences drawn from "limited" evidence. Had they had access to all of the relevant facts, I believe they would have known that there were no honorary coauthors of Darsee's Harvard papers. I developed the overall research strategy — a continuation of a research programme begun in 1968, wrote the grant proposals which supported much of the work, participated actively in the design of the protocols, reviewed the results on a frequent, usually biweekly basis and participated actively in the interpretation of the data. This amply satisfies stringent requirements for coauthorship. Indeed, failure to designate a collaborator with this level of involvement as a coauthor would be misleading. (Had I not been included as a coauthor, Stewart and Feder would have been justified in criticizing me for not shouldering my share of the responsibility.) How is the cause of science served by making such unfair accusations?

Randomization

Stewart and Feder's discussion of the manner in which control dogs were selected is predicated on a misunderstand-

ing. The laboratory had a policy for randomizing control dogs, as follows: when a series of interventions were studied on myocardial infarct size in dogs subjected to coronary occlusion, a comparison was made between infarct size in treated and control dogs. The latter were selected in the following manner: with the first intervention, all control dogs were selected at random and were studied *de novo* and contemporaneously with the treated dogs. With subsequent interventions, when the research fellow was the same and the experimental conditions similar, some control dogs were studied *de novo*, that is, contemporaneously with the treated dogs, whereas the remainder were selected, again at random, from among the controls which had been obtained at random by the same research fellow a short time before.

This practice saved unnecessary sacrifice of animals, but allowed each intervention to be compared to a valid control group. As we advised the NHLBI panel in 1982, it would have been preferable to present this policy explicitly in the papers, but this would not have influenced the results or conclusions. Like all experimental procedures, this ultimately relied on the honesty and integrity of the individual researchers. What did influence the results is that Darsee violated this policy.

Deeply troubling is that Stewart and Feder quote selectively my statements concerning the laboratory's policy of using control dogs. By specifically excluding from the quotation references to key aspects of laboratory policy, they permit the reader to be misled that the Harvard coauthors knew and sanctioned Darsee's dishonest practices in this regard.

Republication

Stewart and Feder allege four instances of "unacknowledged republication" of data from Harvard, one involving two research papers. The first described successful experiments on 21 dogs which were studied 90 minutes and 72 hours following reperfusion after brief coronary occlusion; 14 of these 21 dogs were sacrificed 72 hours after reperfusion and heart tissue was analysed for histology, biochemistry and electron microscopy with the results reported in this paper. Physiological data obtained 72 hours after reperfusion in seven dogs were also reported in this paper; these seven dogs were not sacrificed. It was deemed necessary to apprise the scientific community of the findings, which had important clinical implications. Therefore, we decided to publish these results as the work continued and additional data became available.

The second paper included substantial new information, including the results of successful experiments on 20 additional dogs. That paper reported biochemical, electronmicroscopic and histochemical data on these 20 additional dogs, which were studied either 7 or 14 days following reperfusion. The second paper also reported for the first time: (1) physiological data obtained[4] and 14 days following reperfusion in the 7 dogs for which data obtained 72 hours after reperfusion had been described in the first paper and which were not sacrificed and (2) histochemical data on dogs sacrificed at 72 hours. The second paper recapitulates the earlier findings, making express reference to the first paper twice in the opening paragraph. Would we have done this if we had wished to conceal republication?

Stewart and Feder suggest that there were lapses in standards in the submission of certain Harvard abstracts. However, these abstracts were prepared in strict conformity with the rules of the sponsoring societies. None was submitted for duplicate presentation at national meetings. To suggest that there is something improper in revising an abstract to enhance its quality and clarity is baffling. For example, Abstract 70 was submitted but not accepted for presentation to the American Federation for Clinical Research. After the abstract was revised to improve its clarity and focus, it was accepted as Abstract 78 for presentation seven months later at the scientific session of the American Heart Association. Stewart and Feder's contention that the Harvard coauthors somehow engaged in a questionable practice when they revised a previously rejected abstract before resubmitting that work flies in the face of commonly accepted practice.

Conclusions

Just as it is incumbent on scientists to be meticulous about the manner in which they carry out their work, so it is incumbent on analysts of the conduct of science to be meticulous about their activities. Just as it is damaging to science to make exaggerated claims not supported by data, so is it damaging to science to make exaggerated criticisms and allegations about the conduct of science which are not fully supported by facts. Just as scientists must report the results of all relevant experiments and not select just those which support their conclusions while deleting those which do not, so should analysts of science adhere to the same rules.

Despite my differences with Stewart and Feder's analysis regarding Darsee's work at Harvard, the larger problem which they address, albeit obliquely, namely that of fraud in science, is regrettably still with us and must not be swept under the rug. When we discovered and ultimately exposed all of Darsee's misconduct we were staggered at its extent. Nothing in my more than 25 years in research prepared me to deal with it. In retrospect, there are aspects of the matter that could have been handled better. Perhaps one positive outcome of the unfortunate Darsee affair is that it has helped to focus attention on fraud in science, in part prompting both research institutions and granting agencies to develop more formal rules and procedures for dealing with suspected fraud.

I have thought much about fraud in science during the past five and a half years. Although I have not found any simple, new solutions, a number of issues are clearer to me now.

First, it must be understood by all that to be a scientist is a privilege and that society invests a special trust in all scientists. Whenever that trust is abused it diminishes all scientists. Second, because the costs of serious fraud are so high, we must be more vigilant and more sceptical, always alert to the possibility, however remote, that data we have been shown have been misrepresented. Third, in awarding promotions and grants, undue weight is sometimes given to the quantity as opposed to the quality of a scientist's contributions. This unfortunate practice may lower the quality of science, but I do not believe that it is the root cause of fraud. Fourth, it is usually not possible to explain the deeper motivation for the commission of serious scientific fraud within the framework of normal behaviour. I believe that many instances of scientific fraud represent a form of unconscious self-destructive behaviour which may have aggressive components directed against the perpetrator's supervisors, colleagues and institution; indeed, it mocks science as a whole. Fifth, once a single act of outright dishonesty is discovered, the burden should shift to proving that other data presented by the scientist in question were obtained honestly. Sixth, scientific fraud is a crime and must be investigated like any other crime. Not surprisingly, scientists are not trained to be criminal investigators and few are good at it. Seventh, I believe that laboratory directors and mentors are particularly ill-suited to carry out investigations on their own staff, as they have difficulty shedding their roles as teachers of and advocates for their trainees in order to assume the adversarial role demanded of investigators of a possible crime. Eighth, conditions in laboratories can almost always be improved. It is helpful for laboratories to prepare, adopt and adhere to specific written procedures on the conduct of experiments and the reporting of data.

Finally, I do not believe that the environment in which science is carried out is fundamentally flawed, because the overwhelming number of scientists are honest. Society at large is now grappling with the problem of developing a response to terrorism and with the question of how much of our personal freedom we must give up in an effort to minimize it. In certain respects, fraud poses a similarly vexing problem for science. The analogous question in science is how much are we prepared to change the manner in which we live our scientific lives, an integral part of which involves a sense of trust in colleagues, and how much would we lose in making this change in an effort to minimize (it is probably not possible to abolish) scientific fraud? □

Eugene Braunwald, MD is the Hersey Professor of the Theory and Practice of Medicine at the Harvard Medical School, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA.

APPENDIX 3

PRESS COMMENTS ON STEWART AND FEDER PAPER

(213)

SCIENCE

9 JANUARY 1987
VOLUME 235
NUMBER 4785

Fraud in Science

Fraud in scientific research is unacceptable and inevitable. It is unacceptable because the entire procedure of publishing and advancing knowledge is based on trust—that the literature reports accurate measurements of actual experiments. If each researcher had to go back and repeat the literature, the enormously productive rush of modern science would slow to a snail's pace. Even good intentions are not enough. Sloppy experimentation and poor scholarship are condemned. Outright fraud is intolerable.

Nevertheless, some fraud will exist as long as human beings are doing the experiments. Any system in which advancement, fame, and fortune await a successful practitioner will tempt a certain number of individuals to cut corners. That number may well be smaller in science than in other fields, not because scientists are more moral than others, but because the cumulative nature of science means inevitable exposure, usually in a rather short time.

An oversimplified admonition might be, "You may escape detection by falsifying an insignificant finding, but there will be no reward. You may falsify an important finding, but then it will surely form the basis for subsequent experiments and become exposed." Therefore, there is little percentage in falsifying science, and the speed with which recent examples of this unfortunate human frailty have been revealed is an indication of the pace of modern science. Some newspaper reporters have used recent fraud cases to imply that the structure of science is crumbling or that there is a cover-up, forgetting that the extent of the scientific enterprise has grown a thousandfold since the 1800s. We would expect a greater number of cases of fraud today, but there is no evidence of an increased percentage. And there is no modern equivalent of Piltdown man, a fraud that took years to uncover. Still, it is important that scientists be ever vigilant, and the rash of recent frauds does suggest some dangers in modern science.

One danger arises from the nature of interdisciplinary research. Many papers have numerous authors: investigators in a laboratory that has cloning expertise collaborate with others in a laboratory that has expertise in physical instrumentation and another laboratory that does animal tests to publish a joint paper. The results of this kind of collaboration have had spectacular success, in the main, and no one would wish to limit such joint efforts. Yet when no one person has expertise in all aspects of the research, there can be dangers. A second problem arises when busy scientists, who have too many projects and too little time, supervise projects in which they have infrequent contact with those doing the experiments. Finally, the competitive world of modern science fosters some entrepreneurs who are so intent on the next grant or the big success that they forget that every good experimenter must be his own devil's advocate. A principal investigator must not only devise critical tests for his findings, but must also generate an atmosphere that encourages co-workers to report the bad news as well as the good news.

The procedures recently established by the National Institutes of Health and various universities to deal with fraud seem admirable and appropriate. The punishments for offenders are severe: usually, total derailment of a career. Because the repercussions associated with fraud are so serious, some investigations of such charges take long periods of time, but fairness to the accused is essential. Once guilt is ascertained, the loss of a career in science seems appropriate in many cases. Restitution in some form for the wasted time of those who based further experiments on the false report might be considered appropriate as well. The larger the group, the more interdisciplinary the research, the more competitive the area, the more is the need for watchful skepticism.

Having acknowledged that, we must recognize that 99.9999 percent of reports are accurate and truthful, often in rapidly advancing frontiers where data are hard to collect. There is no evidence that the small number of cases that have surfaced require a fundamental change in procedures that have produced so much good science. To continue the great advances that are being made, we must accept that perfect behavior is a desirable but unattainable goal. Vigilance? Yes. Timidity? No.—DANIEL E. KOSHLAND, JR.

nature

NATURE VOL. 325 15 JANUARY 1987 181

Fraud, libel and the literature

A much-publicized attempt to show that the scientific literature is flawed is not itself above reproach. What science needs is less, not more, solemn attitudes to professional publications.

THE remarkable document on pp 207–214 of this issue should be read only in conjunction with the comment thereon by Dr Eugene Braunwald (pp 215–6) and the following explanation of *Nature*'s role and interest in this unusual affair. Drs Walter Stewart and Ned Feder are two tenured members of the staff of the US National Institutes of Health (NIH) who have carried out an analysis of the research reports and other documents published in the scientific literature by John C. Darsee, who was discovered in May 1981, while a postdoctoral fellow at the Harvard Medical School, to have been guilty of deliberately fabricating data. So much was quickly established by committees of investigation set up by Harvard, by Emory University (where Darsee had previously worked) and by NIH. Stewart and Feder have attempted, by the scrutiny of Darsee's prolific published works, to identify errors, inconsistencies and other flaws that are unconnected with the known falsification of data. The essence of their conclusion is that errors, necessarily subjective in definition, are abundant; this, they say, reflects on the care with which Darsee's coauthors discharged their responsibilities, as well as on the management of scientific journals by their editors and referees.

As will be seen from Braunwald's comment, the methods and the conclusions of this unusual study are disputable; no doubt there will be other responses of the same kind to some of the other illustrations in Stewart and Feder's short list of examples. One obvious difficulty is that, by sticking woodenly to the decision that they would base their study only on the published literature, not asking those responsible for explanations, Stewart and Feder have no doubt been led to count as errors some occasions when authors have merely failed to make their intentions clear.

Yet the document will inevitably raise in readers' minds important questions about the reliability of the scientific literature and of the role of the process of publication in the practice of science. Is the literature in general as flawed as Stewart and Feder claim their "sample of convenience" to be? And, if so, why, and what can be done about it? Many will say that the first assertion has not been proved, so that the second question does not arise, but that is too complacent an opinion. There are also important questions about the pressure to publish that need attention, independently of the analytical part of what Stewart and Feder have written. So why not consider those issues in their own right, without publishing a document whose accuracy is disputed?

Notoriety

The truth is that the document by Stewart and Feder has by now acquired a notoriety comparable with the Darsee affair itself. An early version was submitted for publication in *Nature* in 1983; two reviewers (outside the United States) took the view that the manuscript, while of absorbing interest, would require that certain editorial questions of principle should be decided before publication could be seriously considered. The paper was accordingly sent for comment to a dozen or so of the coauthors identifiable even from the anonymous list of Darsee's publications in the accompanying bibliography, three of whom replied through their attorneys with the opinion that publication would be defamatory, an opinion already given by *Nature*'s own legal advisers.

Litigious readers may note, at this stage, that, at least under British libel law, it could well be libellous to publish an innocuous version of a manuscript accompanied by the intelligence that earlier versions had been defamatory of still identifiable people. Readers should therefore know that what follows has been discussed with some of those concerned, who have agreed that it should be published so as to bring out into the open an issue of which Stewart and Feder have recently made much.

Replication

Thanks to the marvellous powers of word-processors, the first version of the manuscript was quickly followed by many others. At a meeting in September 1984 with two of the co-authors and their attorneys, the editor of *Nature* was left in no doubt that publication of any of the versions then extant would be considered defamatory, an opinion with which he agreed. At the same meeting, a number of constructive suggestions were made for modifications that might make a publishable manuscript. In the event, the need to transact normal business, and the fearsome volume of comments and counter-comments that had accumulated, meant that it was not until the near the end of 1984 that an edited version of the manuscript which *Nature* considered to be acceptable was being transmitted by telefacsimile from London to the authors at NIH. When only 26 of more than twice as many pages had made the transit, the recipient authors telephoned hotly to say that the changes made the manuscript unacceptably anodyne, and that it would be withdrawn; they were told that *Nature* would be prepared to reconsider a new version of the original amended along the lines suggested if they failed in their expressed intention of having the original published elsewhere. The version now published is derived, mostly by abbreviation, from one submitted a year ago.

These tedious details are relevant to present circumstances because of the way in which Stewart and Feder have injudiciously, during the past year, sought to argue that the threat of libel suits has been used to suppress an important comment on the integrity of the scientific profession. This was the essence of their case to a congressional subcommittee in the United States last year, in a series of statements to the *New York Times* (26 April 1986) and, more recently and more damaging, in a letter (accompanied by an earlier unexpurgated version of the manuscript) asserting that publication had been prevented by the threat of libel suits; that letter was sent to all 1,800 members of the US National Academy of Sciences as well as to many of the authors' correspondents on quite unrelated matters. At a guess, upwards of 2,000 people have been circularized in this way.

Integrity

The issue thus raised has an important bearing on the conduct of the scientific literature, which depends for its integrity on its accessibility and openness; in principle, all scientists are free to offer their contributions. Dissent from accepted theories is allowed and even welcomed. Novel ideas may be published for the

sake of the stimulus they provide, but with no implied guarantee that they are in some absolute sense 'correct'. Moreover, authors are frequently made painfully aware by their experience that publications are hostages to fortune; it frequently occurs that the publication of a set of data and of a plausible hypothesis to account for them will stimulate others to produce further data whose effect is to contradict the hypothesis, to the chagrin of the original authors. Often, publications are found to contain self-vitiating errors, whose authors then have even redder faces. Such honest mistakes are an inescapable part of science.

So why does not this process of cut and thrust, personally hurtful to those concerned, require that scientific journals employ armies of lawyers vetting everything they publish? Because of the principle that, while it is a proper and necessary function of authors to correct what appear to be erroneous data or inferences in the literature in publications of their own, most keep to the convention (which is otherwise enforced by editors) that the identification of an error is not followed by the assertion that the perpetrator is a rogue, let alone by speculation about the motives that have engendered this condition. Stewart and Feder have complained that they could not publish what they chose because of "threats of libel". They have not understood that the unfettered right to publish scientific data does not equate with a right to denigrate others' characters.

For all this, what Stewart and Feder have written deserves close attention. At the very least, they have shown that a group of published papers do not provide a full explanation of the data recorded in them. Even if the "bizarre pedigree" they have found (see p.208) would be less puzzling if the ages given in the original paper were intended to mean something other than chronological age, there is nothing in the paper to suggest that simple readers would differ from Stewart and Feder in their interpretation of the paper. Similarly, while Braunwald's defence of the use of "historical controls" is entirely proper, it remains a fact that the procedure that should have been followed (Darsee cut corners instead) is not adequately explained.

Imprecision

How can imprecisions of this kind have found their way into published papers? Stewart and Feder are unforgiving in their opinion that Darsee's coauthors should have spotted them, and have got rid of them, but that asks a lot of flesh and blood. If the documents concerned were drafted in the first place by Darsee, often with the intention of rendering plausible a set of data derived from thin air by fabrication, it requires very little imagination to appreciate the difficulty of rooting out minor errors at source; those who spend time on committees know how much initiative rests with the one who writes the minutes. Stewart and Feder themselves appear to have spent a long time, assisted by the hindsight that Darsee was a fraud, scrutinizing just over a hundred publications. Braunwald also makes a cogent point in saying that the detection of the major error, the fraud, would have required the repetition of the experiments, defeating the objective of research by teams of people. But the general conclusion at which Stewart and Feder arrive, that coauthors should be more zealous in their scrutiny of what is about to be published in their names, anodyne though it may be, cannot be a bad prescription.

Coauthorship as such is a necessary consequence of the way that science is now practised, enabling people with different skills to work together on a single project; it is also capable of being abused. Stewart and Feder define as "honorary" those coauthors who have not been directly involved in the research reported in published papers. This is a restrictive definition, especially where the members of a research team include senior and less senior people. Most people, probably including Stewart and Feder, would accept the Braunwald position that a person whose role has been to conceive of a research programme, to raise the funds with which it is supported and to supervise progress, is legitimately a coauthor of the research. This is the standard relationship between academic researchers and their graduate students, for example.

On more general grounds, Stewart and Feder are right to insist that the literature is corrupted if its users cannot be assured that all the authors of a publication have played some essential part in its development. There is evidence, mostly anecdotal, quite unconnected with the Darsee case, to show that what they call "honorary coauthorship" is commonplace. In molecular genetics, for example, those who supply clones of some genetic material for use by others are often listed among the authors of the ensuing research. Some coauthorship is political in flavour, as when the prime movers in international projects cite sleeping partners as authors for the sake of continued opportunities to investigate otherwise inaccessible phenomena. The practice that the head of a laboratory should as a matter of course appear as an author of all its publications, still accepted as a convention in some Indian laboratories, for example, is rapidly on the way out and should be helped on its way by Stewart and Feder's article, which will remind those concerned of the risks they run. Some journals, responding to the *samizdat* publication of that article, have gone so far as to decide that future coauthors will be required to declare their full involvement in work sent for publication (a practice, interestingly, followed as a matter of course by groups of authors from the Soviet Union). But by itself, such devices will not end the abuse of coauthorship.

Publications

How common are errors of the kind that Stewart and Feder categorize? And how serious are they? Their own study, some of whose conclusions are disputed, concerns an exceptional sample of the literature. But, shocking though it may seem, the experience of those concerned with the management of the literature is that errors of all the kinds listed are far from being rare. Even when authors do their best, honest mistakes crop up repeatedly. (Scientists are too often indifferent to the plain truth that awkward or ambiguous phraseology is one of the most common reasons why scientific publications are misunderstood.) Luckily, as the world knows, most of these difficulties are cleared up by the informal correspondence between interested parties which is now as much a part of the process of science as the formal literature. More serious faults are less easily dealt with. Instalment publication is often used to magnify a laboratory's output and, as all journal editors know, there are also many occasions when authors are less than diligent in referring to the intellectual sources of their work. Ignorance is often part of the explanation, but ignorance can cloak malign intent.

None of this bears directly on fraud, or on the prevention of it, but there are good reasons for believing that the ordinary defects of the literature arise from the undue importance now attached to the process of publication, largely because of the way in which the careers of younger scientists and the continued success even of successful laboratories have come to depend on what is sometimes known as 'output', which is another name for the number of research articles published, perhaps weighted by length and by the status of the journals in which they have appeared. Some assessment of this kind is the basis of what is called peer-review in the United States (strictly, not to be confused with the process by which journals base their decisions on what to publish on the opinions of external referees). As the economists would say, the research community has externalized the problem of making judgements about the qualities of its own members by brooding on bibliographies. Everybody's interest is to find a less slovenly way of making intellectual judgements about people and institutions. The recipe implicit in the Stewart and Feder argument, that zealous suspicion of everybody within sight is the way to ensure the integrity of the scientific literature, is of course a recipe for disaster, a road to general mistrust, a licence for every would-be whistle-blower and a means by which the literature would be made yet more solemn. Decoupling the literature from promotion prospects would be a better goal. □

News & Comment

Integrity of Research Papers Questioned

Two researchers from NIH allege that the scientific literature may be full of minor errors; others argue they have very little evidence

In the spring of 1983, Walter W. Stewart and Ned Feder "set aside" their usual laboratory work at the National Institutes of Health to embark on a wholly unusual investigation of the validity of the scientific literature. Through that they hoped to shed light on the question of whether coauthors of scientific papers, journal editors, and referees are "meeting the standards conventionally accepted as necessary" to assure that the literature is accurate.

Stewart and Feder were interested in a number of points including the factual consistency and accuracy of papers, duplicate or "republication" of data in abstracts and articles, and the appropriate assignment of credit and responsibility when it comes to deciding whose name will go on a manuscript and whose will be left off. They took as their study sample 109 papers written by John Darsee and 47 coauthors, some from Emory University School of Medicine, others from the Harvard Medical School laboratories of cardiologist Eugene Braunwald. In 1981, Darsee was discovered to have committed fraud at Harvard and it was later discovered that he also committed fraud in medical school at Emory, and even before that as an undergraduate at Notre Dame.

Using the Darsee papers as a "sample of convenience," Stewart and Feder concluded that those papers are full of errors and inconsistencies, quite apart from the frankly fraudulent data, and that the coauthors could have spotted those errors had greater care been exercised in preparing the manuscripts for publication. Stewart and Feder also assert that journal editors and referees failed to attend sufficiently to detail to catch what the NIH scientists regard as obvious errors.

A manuscript reporting their observations was first submitted to *Nature* in the fall of 1983, but was subsequently withdrawn when the authors concluded that changes mandated by legal considerations of libel watered it down too much. Stewart and Feder went on to offer their paper to several journals, including *Science* and the *New England Journal of Medicine* where it was rejected on editorial grounds. In a letter that has become part of the public record through Freedom of Information Act requests for access to legal and editorial documents on the Stewart-Feder manuscript, *Science* editor Daniel E. Koshland, Jr., wrote them in July 1985 and said, "I do not concur with you that details of this case are extraordinarily important. . . ." *New England Journal* editor Arnold S. Relman rejected the version he saw. "What was true wasn't new, and what was new wasn't true, or in any event very misleading," he recently told *Science*. But other editors told Stewart and Feder they thought their paper was important and in the public interest, but too expensive to handle because of its potentially libelous nature.

Almost everyone agrees that coauthors could take their responsibility for the accuracy of data more seriously.

In addition to submitting their manuscript to many journals, Stewart and Feder also reported the essence of their findings to a committee of the Congress on 26 February 1986, when they asserted that fear of libel was preventing publication. And, within the past few months, they sent copies unsolicited to every member of the National Academy of Sciences. Because this study had NIH approval, NIH paid copying and postal costs.

Finally, the saga of the Stewart-Feder manuscript now comes full circle with publication of a much edited version in the 15 January issue of *Nature*. In an accompanying editorial, John Maddox calls attention to a companion article in which Harvard's Braunwald offers a rebuttal to allegations that he and his colleagues allowed sloppy work to be reported in the scientific literature. (Researchers at Emory have yet to be heard from.)

In his comment in *Nature*, Braunwald also recalls, "When we discovered and ultimately exposed all of Darsee's misconduct we were staggered at its extent. . . . Perhaps one positive outcome of the unfortunate Darsee affair is that it has helped focus attention on fraud in science, in part prompting both research institutions and granting agencies to develop more formal rules and procedures for dealing with suspected fraud." (The Association of American Medical Colleges and the Association of American Universities each produced guidelines. Formal procedures were recently established by NIH.)

One of the more unusual features of this event is the suggestion, implicit in the paper and its publication, that one can generalize from the faults in the Darsee papers to the scientific literature generally. If these papers are full of errors, is it not possible that the scientific literature is chockablock with papers whose data may be inaccurate? To this, almost everyone answers—"You cannot say."

Stewart and Feder take this on themselves. In their *Nature* paper, "The integrity of the scientific literature," they write, " The results of our study naturally raise the question: what fraction of papers in the biomedical sciences are not supported by primary data at the time of publication? The question cannot be answered with the papers in our sample, since most are known to be partial or complete forgeries and thus are atypical." Further, Stewart and Feder make this observation about the nature of many of the problems they identified in that sample: "We emphasize again that many of the errors and discrepancies are minor, for example a small discrepancy between a numerical value printed in a table and what is supposed to be the same value printed in the text."

Why, then, publish it at all. In his editorial, *Nature*'s Maddox says, "The truth is that the document by Stewart and Feder has by now acquired a notoriety comparable with the Darsee affair itself."* Maddox goes on to say, "Stewart and Feder have complained that they could not publish what they chose because of 'threats of libel.' They have not understood that the unfettered right to publish scientific data does not equate with a

*The Darsee affair was widely reported in the press, including *Science*, which published the following News and Comment articles: 29 January 1982, pp. 478–482; 12 February 1982, pp. 874–876; 1 April 1983, pp. 31–35; 27 May 1983, p. 936.

right to denigrate others' characters."

Nevertheless, the Stewart-Feder paper focuses attention on questions about scientific publication that are widely regarded as important. Maddox writes that "the general conclusion at which Stewart and Feder arrive, that coauthors should be more zealous in their scrutiny of what is about to be published in their names, anodyne though it may be, cannot be a bad prescription."

Even if many of the errors Stewart and Feder identify can be explained, the fact remains that the errors are there. In one instance, which they describe as "so glaring as to offend common sense," Stewart and Feder call attention to a family pedigree of individuals with heart disease that shows one man having fathered four children ages 8, 7, 5, and 4 by the time he was only 17 years old. That the ages noted on the pedigree were *meant* to indicate the age at which heart disease was diagnosed is of little help to readers of the original article who would have no way of knowing that.

An interesting aspect of the Stewart-Feder story is the NIH's official role in all this. As intramural scientists, the two researchers enjoy substantial freedom to tackle whatever projects they think are important, and at first there was no objection to their undertaking an analysis of the literature along with their laboratory research. But as the Darsee study came to occupy more and more time, questions arose within the National Institute of Diabetes and Digestive and Kidney Diseases, where they work, about the fact that their laboratory science had all but come to a stop.

The matter was taken to the office of the NIH director where, in the spring of 1984, Joseph E. Rall, deputy director for intramural research, and others made an affirmative decision that Stewart and Feder should be permitted to complete their literature study. William Raub, newly named deputy director of NIH, was in on the decision as head of the NIH office on research fraud. "We agreed that it was a proper scientific inquiry," Raub told *Science*, adding that no one envisioned that completion of the project would occupy Stewart and Feder full time for two-and-a-half years more. Nor did they appreciate the amount of time the NIH's own lawyer would spend helping them produce a legally defensible manuscript.

Only a couple of the coauthors retained legal counsel to block publication of the original versions of the Stewart-Feder paper, but it only takes one to raise roadblocks that can cause interminable delay unless allegations are provable absolutley. Bancroft Littlefield of Boston took the lead as Braunwald's counsel.

In the end, where earlier versions and congressional testimony talked about "serious misconduct" and "statements for which the evidence indicates that they knew or should have known of their falsity," the published paper uses less damaging language, referring instead to "incomplete" rather than "misleading" statements, for instance. Stewart and Feder are not entirely happy with the changes and *Nature* notes that some editorial changes (including the latter) were made without their consent. But any threat of suit seems to have been averted, at least from the Harvard coauthors, and Littlefield says it is better now that the paper be published than to allow further charges that Braunwald and others are using legal threats to stiffle the free flow of information.

What now? Stewart and Feder will face demands from their institute that they return to laboratory science, and they have said they want to do so. Their own work has been with a class of agents known as "Lucifer dyes" that enable visualization of the branching patterns of nerve cells. Feder, 59, graduated from Harvard College and Harvard Medical School. Stewart, 42, graduated from Harvard College but lacks a doctorate. The two have worked together at NIH for many years. In 1981 Stewart published two highly regarded and cited research papers on the dyes.

But they have also expressed an interest in continuing to work on questions about the integrity of scientific literature to answer questions that the study of the Darsee papers could not answer. The appropriate balance between laboratory research and other professional inquiry by intramural scientists is an issue that NIH officials will have to grapple with.

At the conclusion of their *Nature* paper, Stewart and Feder suggest auditing a random sample of the literature to check on its integrity. But they also note that there are potentially serious costs to "examining the practices of scientists" this way. "Systematic examination of scientific practices might even weaken the fabric of trust that is essential to the functioning of science," they write, adding that even though science is vulnerable to abuse, it is "perhaps even more vulnerable to harm by regulation." ■

BARBARA J. CULLITON

Fraud, Laxity in Research Are Detailed

Controversial Report Questions Effectiveness of Scientists' Self-Regulation

27 Jan 87

By Boyce Rensberger
Washington Post Staff Writer

Two National Institutes of Health scientists have suggested that carelessness in gathering and reporting scientific data may be more common than is usually supposed.

The NIH researchers based their conclusions on a study of 109 scientific publications coauthored by Dr. John Darsee, the Harvard medical researcher who was discovered in 1981 to have faked scientific findings, and 47 of his colleagues.

The NIH scientists, Walter W. Stewart and Ned Feder, say the papers contain so many obvious discrepancies that Darsee's colleagues, all at either Harvard or Emory University, who lent their names as coauthors either knew or should have known that there were problems with the papers.

None of the colleagues has been implicated in any scientific misconduct and the study published in the Jan. 15 issue of the British journal Nature does not name them.

It does suggest that fraudulent scientific activity sometimes continues undetected because scientists, even though not personally committing fraud, are sometimes too casual in lending their names and scientific prestige to the publications of data forged by colleagues.

The study, completed more than three years ago, is one of the most controversial in recent scientific history. Several professional journals refused to publish it on grounds that it was libelous. Lawyers for some of Darsee's coauthors bombarded the journals and the NIH scientists with threatening letters.

Last February, when it appeared the threats might prevent the study from ever being published, the controversy became the subject of hearings before a House Judiciary subcommittee looking into the "chilling effects" of libel threats on the willingness of editors to publish controversial articles.

Now, however, the report, after numerous rewrites and the elimination of some entire sections, has been published in what the authors called "an expurgated version." Despite that, they say they are happy to see it published.

The story of the article and the difficulty it encountered illustrates the sensitivity of the scientific community toward some forms of self-criticism. Criticism of theories and research methods on objective grounds is widely accepted as essential to the quest for scientific truth. Criticism of the more subjective aspects of professional conduct, often based on ethical grounds, however, is another matter.

Stewart and Feder, the study's authors, are career scientists at the National Institute of Arthritis, Diabetes and Digestive and Kidney Diseases, a part of the National Institutes of Health. While conceding that they did not study a representative sample of the scientific community as a whole, they said their findings raise questions about the way that community, almost entirely self-regulated, is able to detect and deal with professional misconduct.

"Self-regulation is a privilege that must be exercised vigorously and wisely, or it may be lost," Stewart and Feder wrote in their report.

The form in which the report was published illustrates the professional sensitivity. At the end of the article, for example, is an unusual editor's note: "Some editorial changes have been made in this manuscript without the consent of the authors."

Nature editor John Maddox also added a two-page editorial criticizing the report as "disputable" and "not itself above reproach," but explaining that he published it anyway because it focuses on legitimately questionable practices in science.

Also accompanying the report is a rebuttal by one of Darsee's more prominent coauthors, Dr. Eugene Braunwald of Harvard Medical School, whose lawyer tried to block publication of early drafts of the report. Braunwald acknowledges that the report raises troubling questions but argues that most faults cited in the Darsee papers are minor and reflect no fundamental flaws in the practice of science.

Stewart and Feder's study consisted of examining all the papers Darsee coauthored to find errors that they believed alert coauthors should have spotted.

In the 18 publications that were full-length research papers, Stewart and Feder counted an average of 12 errors and discrepancies per paper and a maximum 39 errors in one paper. They were of a type that Stewart and Feder say should have raised the coauthors' suspicions about the data, most of which was provided by Darsee and is now known to have been faked.

For example, one paper subsequently withdrawn as fraudulent included a family tree in which a 17-year-old male is said to have had four children from 4 to 8 years old. Because the age is given several times in the paper, it was deemed "unlikely to have been a misprint. The boy would therefore have been 8 or 9 years old when he impregnated the mother of his first child.

"This bizarre feature of the pedigree, which perhaps could have raised questions about the validity of the entire paper, was apparently not noticed by coauthors or referees [independent reviewers of the article before acceptance by The New England Journal of Medicine]," Stewart and Feder wrote.

The same paper contains discrepancies between lab results given in the text and also presented in a table, and contradictory statements to the effect that all the patients studied were alive at the time of the study and, in another place, that three had died before the study.

Stewart and Feder cite many examples of similar discrepancies and implausibilities from other papers and conclude that "the abundance of errors that we observed in our sample obviously interferes with the accuracy of the scientific literature. Every scientist has, at a minimum, an obligation to ensure that what is published under his name is accurate."

In some cases, Stewart said, the report's original interpretation of findings was "watered down" by Maddox's editing. Where an earlier draft, for example, said that Darsee's coauthors "knew or should have known of the statements that were false," the wording was changed to say they "could have known of the statements that were imprecisely misleading."

"How can imprecisions of this kind have found their way into published papers?" Maddox asked in his editorial. "Stewart and Feder are unforgiving in their opinion that Darsee's coauthors should have spotted them, and have got rid of them, but that asks a lot of flesh and blood."

Maddox concludes, however, that "coauthors should be more zealous in their scrutiny of what is about to be published in their names."

At the heart of the problem, Stewart and Feder say, are the traditional practices of coauthorship.

Modern scientific research is often a collaboration of several scientists. Sometimes they do the work jointly; sometimes there is a division of labor according to each scientist's specialty. In the latter case, each scientist normally trusts the others to be honest. In some laboratories, it is the custom to include as coauthors persons who were lab chiefs offering encouragement but who did none of the research or persons who lent equipment or who contributed funds.

In some cases, Stewart and Feder say, the basis of authorship is so tenuous that the scientists who lend their names are merely "honorary authors."

"Honorary authorships falsify the assignment of responsibility for published research and increase the likelihood that inaccurate data will be published," they wrote. "The honorary author is in a poor position to judge the validity of the work, yet he often lends a prestige that may lull other coauthors, the reviewers or the readers into uncritical and inappropriate acceptance."

Although none of Darsee's coauthors is named in the report, Nature asked one, Braunwald, to respond.

While agreeing that the paper raises issues that "are important and deserve attention and thoughtful analysis," Braunwald wrote that the Stewart and Feder paper contained errors and questionable interpretations of its own and that it unfairly tarnished Darsee's coauthors through "guilt by association."

Braunwald rebutted several alleged defects in papers that even Stewart and Feder said were minor, but did not address some of the major discrepancies. He also rejected the allegation that any of the Harvard coauthors were merely honorary. All, he said, played substantive roles in the research.

Braunwald suggested the Stewart and Feder report could be as damaging to science as the defects it purports to address if it led to scientists assuming the adversarial roles of watchdogs.

"How much," he wrote, "are we prepared to change the manner in which we live our scientific lives, an integral part of which involves a sense of trust in colleagues, and how much would we lose in making this change in an effort to minimize (it is probably not possible to abolish) scientific fraud?"

New Scientist 22 January 1987

Researchers roll back the frontiers of fraud

Andy Coghlan

SCIENTIFIC fraud and carelessness among researchers could be widespread, warns a study in last week's issue of *Nature*.

After an investigation of America's biggest scientific scandal in recent years, Walter Stewart and Ned Feder from the US's National Institutes of Health in Bethesda, Maryland, conclude that "lapses from accepted standards may be more common among biomedical scientists than is currently appreciated". They call for more studies "to determine whether our results are representative of other groups of scientists".

Feder and Stewart have spent the past five years examining a scandal surrounding the work of John Darsee, once hailed as "one of the most remarkable men in American medicine".

Darsee spent 13 years working at two of the US's leading centres for research into heart disease—Emory University School of Medicine and Harvard Medical School. He admitted in May 1981 to a single case of fabricating data. Later investigations by colleagues and official committees revealed that more than 100 of his publications were forgeries.

Feder and Stewart began their study in 1983, two years after the scandal emerged. Their report focuses on whether the 47 coauthors of Darsee's forgeries knew of—or should have noticed—the forged data.

Only 12 of the 47 are cleared by Feder and Stewart of lapses from accepted standards. They argue that coauthors had been careless when checking Darsee's work. "Some errors are so glaring as to offend common sense," they say.

In one of Darsee's 18 full papers, a 17-year-old male is reported to be the father of four children, aged 8, 7, 5 and 4. "Thus the father was 8 or 9 when he impregnated the mother of his first child and 9 or 10 when he impregnated the mother of the second," say Feder and Stewart. "This bizarre feature . . . which perhaps could have raised questions about the validity of the entire paper, was apparently not noticed by coauthors or referees," they add.

Feder and Stewart blame the coauthors for a variety of sloppy practices. They accuse them of failing to retain important experimental data, of failing to acknowledge when control data from previous publications were reused and of misreporting experimental practices.

Camera Press

Darsee's forgeries sullied Harvard's hallowed reputation

Though Feder and Gilbert admit that outright fraud is "presumably rare", they warn that the abundance of errors and lapses observed debase the scientific literature. They blame a combination of factors for slipping standards, including extreme competition for research grants and the absence of an accepted code of conduct for scientists.

The two investigators draw attention to the "extraordinary emphasis" commonly placed in the American scientific community on the "sheer quantity of publications".

They recommend further study of the prevalence of fraud and carelessness in research, by reviewing a random sample of published papers. But they warn against heavy-handed policing. It could "weaken the fabric of trust essential to the functioning of science", they say.

"Scientists have to an unusual degree been entrusted with the regulation of their own professional activities. Self-regulation is a privilege that must be exercised vigourously and wisely, or it may be lost," they warn.

Feder and Stewart's own scientific integrity comes under fire in the same issue of *Nature*. Eugene Braunwald, defending the 23 Harvard scientists who co-wrote Darsee's work, accuses Feder and Stewart of inaccuracies and of subjectivity.

"The paper repeatedly and unfairly connects Darsee's fraud at Harvard to his coauthors there through a process of guilt by association . . . Many of [their] allegations are misattributions of Darsee's wrong doing to his coauthors," says Braunwald, a professor at Harvard Medical School's Brigham and Women's Hospital in Boston, Massachusetts.

But he admits that fraud is "regretably still with us and must not be swept under the rug. Society invests a special trust in all scientists. Whenever that trust is abused it diminishes all scientists". □

The Editorial Notebook

Fraud and Garbage in Science

29 Jan 87

The Present System Of Quality Controls Screens Out Neither

An elaborate case of scientific fraud, first discovered at the Harvard Medical School in 1981, has turned out to have another, more surprising dimension. Besides fraud, according to an analysis in Nature magazine, the papers of Dr. John Darsee and his co-authors were studded with obvious errors and inconsistencies. Yet even these failed to alert the peer committees, editors and referees who vet the quality of medical research.

Science has its share of cheats and corner-cutters. But scientific frauds are especially disturbing because many sail past standard quality control procedures and are only detected accidentally. If the quality controls aren't detecting even outright fraud, how well do they sift out garbage? Without quality assurance, can new research be relied on?

The Piltdown skull, discovered in 1912, is science's best known fraud. The motive, perhaps, was to prove that the earliest man was English, or to make monkeys of the Englishmen who hailed the skull's antiquity over its German rivals. Modern frauds usually have a narrower goal — inflating the forger's list of scientific publications. Published articles, the coin of success in science, weigh heavily in securing both university jobs and Federal grants. Since few appraisers have time to read the articles, quantity can easily outweigh quality.

John Darsee spent his research career at Emory and Harvard Universities, two leading cardiology centers. He was a prodigious worker, with 109 scientific publications to his credit. He generously spread the credit among 47 fellow scientists who signed their names to his papers.

The quality control system of medical research gave him no problem; he breezed through it every time. His downfall was his success: colleagues suspicious of his productivity caught him inventing data on a heart drug, and he was eventually dismissed.

There was more to his output than just fraud. Almost all of his 18 major research articles contain errors — an average of 12 per paper — that are obvious on careful reading, according to Walter Stewart and Ned Feder of the National Institutes of Health. Why weren't these noticed by journal editors and referees, or by Dr. Darsee's 47 co-authors?

Eugene Braunwald, Dr. Darsee's chief supervisor at Harvard, disputes some of the alleged errors and describes others as trivial. It's also true that Drs. Stewart and Feder scrutinized the Darsee papers with the advantage of knowing of his fraud. Still, Dr. Darsee and his co-authors were dumping much valueless research into the scientific literature and the errors, trivial or not, were red flags that amazingly caught no one's eye.

The fraud reflects on Dr. Darsee. The failure to notice it during his 12-year research career mocks the quality assurance procedures in biomedical science. That point is not yet widely understood by the science establishment, which is still prone to blame only the individual miscreant. As Daniel Koshland, editor of Science magazine, wrote this month, "There is no evidence that the small number of cases [of fraud] that have surfaced require a fundamental change in procedures that have produced so much good science."

But the few detected cases of fraud are only the extreme. Careless science and manipulated data are probably far more common and science's quality control system may also be letting in a lot of garbage. Poor research, like most fraud, may do little harm but it costs just as much to produce as good science and crowds out young researchers seeking funds to get started. A breakdown of science's quality assurance, as signaled by the Darsee case and other frauds, sounds an alarm worth more attention.

NICHOLAS WADE

Letter: On Research

Honest Science Errors Aren't 'Garbage'

To the Editor: 7 Feb 87

In his Editorial Notebook of Jan. 29 ("Fraud and Garbage in Science") Nicholas Wade contends that science journal editors and reviewers are unable to screen out fraud or poor quality research (which he terms "garbage"). He is right about fraud but wrong about the "garbage." Furthermore, his discussion betrays a woeful misunderstanding of the review practices used by the journals and fails to perceive important distinctions between fraud, "garbage" and error.

Reputable peer-reviewed science journals generally do a good job in selecting the best work submitted to them, but they probably will not detect an adroitly fraudulent study. That is because in science there must always be a presumption of honesty. Contrary to Mr. Wade's assertion, this presumption is not invalidated simply by the discovery of numerical errors or inconsistencies in a manuscript. To challenge an author's trustworthiness on such bases is to confuse infallibility with integrity and to allow no room for honest error. That would create an atmosphere of intolerable suspicion and distrust and would threaten the viability of the entire research enterprise. Unless the culprit's acts are detected by his own co-workers or supervisors, fraudulent studies are not likely to be recognized until other scientists try to duplicate or extend the work.

There is no evidence that outright fraud taints the work of more than a tiny fraction of biomedical scientists, or that its incidence is increasing. What does seem to be on the rise is the frequency of public reports of fraud — a phenomenon not unexpected in view of the recent growth in the number of biomedical scientists and the intensity of media coverage.

Whatever its incidence, I agree with Mr. Wade that fraud in science is a serious problem requiring more study than heretofore. However, I can't agree there's been a "breakdown" in science's quality assurance system. Journalists ought to avoid overheated rhetoric lest they destroy public confidence in vital institutions that have served society well. Neither biomedical research nor science journals should be exempt from examination or criticism; nor should the subject of "fraud and garbage in science" be considered taboo. I am simply suggesting they deserve a more informed and balanced discussion.

ARNOLD S. RELMAN, M.D.
Editor
New England Journal of Medicine
Boston, Feb. 1, 1987

Study Alleging Scientific Misconduct Is Published; Findings Disputed

A three-year fight over a study of scientific misconduct has finally erupted into print in a scientific journal.

The study, which had been extensively discussed and privately distributed but never before published, appeared in the January 15 issue of the British journal *Nature*, along with a rebuttal and a critical editorial.

Two researchers at the National Institutes of Health, Walter W. Stewart and Ned Feder, said last spring that their study of misconduct and carelessness in scientific research had not been published because some of its subjects had threatened lawsuits against journals that were considering printing it (*The Chronicle*, May 21, 1986).

After the article was rejected for publication, Mr. Stewart and Dr. Feder went before a Congressional committee, talked to lay publications, and sent a letter and a version of the study to all 1,800 members of the National Academy of Sciences.

Editors, Co-Authors Faulted

In the *Nature* article, Mr. Stewart and Dr. Feder said that, in a study of publications written in part by a John C. Darsee, who was found by previous investigators to have falsified data, they had found a lack of vigilance by journal editors and Dr. Darsee's co-authors in detecting errors. They also uncovered other "lapses from generally accepted standards of research," they said.

For the study, the two N.I.H. researchers looked at 109 publications credited either in part or in whole to Dr. Darsee, a cardiology researcher at Emory University and at the Harvard University medical school who was found in 1981 to have forged data.

They said that in the majority of his research papers they had found discrepancies and errors "that can be recognized simply by examining them carefully."

A *Nature* editorial differed with the article's conclusions.

"By sticking woodenly to the decision that they would base their study only on the published literature, not asking those responsible for explanations, Stewart and Feder have no doubt been led to count as errors some occasions when authors have merely failed to make their intentions clear," the editorial said.

Early Study Found Defamatory

The editorial said an early version of the manuscript had been found to be defamatory by *Nature*'s legal advisers and by lawyers representing three of Dr. Darsee's co-authors.

Mr. Stewart and Dr. Feder did not list Dr. Darsee's 47 co-authors, but they have acknowledged that the names could be discovered.

In a rebuttal, one of Dr. Darsee's co-authors and his former supervisor, Eugene Braunwald, professor of the theory and practice of medicine at the Harvard medical school, disputed the accusations of misconduct, and made countercharges that the study's authors were guilty of their own brand of scientific sloppiness.

"Just as it is damaging to science to make exaggerated claims not supported by data, so is it damaging to science to make exaggerated criticisms," Dr. Braunwald wrote.

Both the editorial and Dr. Braunwald's reply reflected on some of the causes and potential cures for scientific misconduct.

Said Dr. Braunwald: "Many instances of scientific fraud represent a form of unconscious self-destructive behaviour which may have aggressive components directed against the perpetrator's supervisors, colleagues and institution; indeed, it mocks science as a whole."

—DAVID L. WHEELER

SCIENCE

Tempests in a Test Tube

Two new studies ask why scientists cheat

It was an innocent enough question: how do scientists behave when no one is looking? But it has produced an incendiary answer: not too well, reports a paper this month in the British journal Nature. Although the foundation of science is trust and its goal is truth, the field has been rocked by scandal dating back at least to Piltdown man, the faked fossil skull discovered in 1912. More recent frauds have triggered a wave of soul-searching over what constitutes fraud and why it occurs. Now the Nature study and a new book bring the disturbing message that the *way* science is practiced, as much as unethical individuals, breeds misconduct—and that it may occur more often than scientists realize.

The litany of frauds in "False Prophets"* by Alexander Kohn of Tel Aviv Medical School makes it seem that ethics holds no greater sway in the lab than it does on Wall Street. Kohn believes that most instances of scientific cheating are eventually uncovered, but his theory of why it occurs implies that fraud could become a growth industry. For one thing, he notes, when a science is undergoing a revolution, as molecular genetics is, it is easier to cheat because the flood of new ideas is too great for scientists to monitor. Also, fraud is more common in biology than in "hard sciences" like physics, where exact laws of nature serve as a template against which results are measured. The many variables in biomedicine, in contrast, allow erroneous data to slip by.

Faked research: In an infamous fraud at Harvard Medical School in 1981, John Darsee, then 33, was testing heart drugs on dogs. His finding that two drugs limited damage after a heart attack contradicted results from other labs, and when challenged Darsee could not substantiate his work. It turned out that he had fabricated data in 109 publications. While condemning Darsee, a National Institutes of Health panel also concluded that lab chief Eugene Braunwald's remoteness from day-to-day lab work helped keep the falsifications from being discovered, a charge Braunwald denies.

*236 pages. Basil Blackwell. $24.95.

RICHARD HOWARD

Do lab practices breed misconduct? *Braunwald*

The role of Darsee's colleagues is also the subject of the Nature paper by Walter Stewart and Ned Feder of NIH. In scrutinizing his publications, they found numerous additional "lapses"—not necessarily deliberate but some "so glaring as to offend common sense," such as a family tree showing a 17-year-old father of four. Though it is not clear whether the errors were typos or sheer fabrications, Stewart and Feder say Darsee's coauthors should have caught them. The NIH team blames the common practice of "honorary authorship," in which illustrious superiors who had little to do with an experiment are listed as coauthors. The practice, writes the NIH team, may lull reviewers into "uncritical and inappropriate acceptance."

Braunwald responds that such accusations are "quibblings." But others agree that some practices common to science can erode its integrity. A government grant system that rewards researchers who churn out the most reports, for instance, promotes honorary authorship and pressures investigators to publish or perish—which may induce outright fraud. In the most celebrated case, in 1974, William Summerlin of Memorial Sloan-Kettering Cancer Center in New York darkened skin on a white mouse with a pen to simulate a graft. He also mistakenly claimed that he grafted skin from a white mouse onto a purebred dark mouse, when it was actually a hybrid. Frauds arise, too, because agencies seldom support scientists who repeat each other's work. As a result, says Arthur Neufeld of Boston's Eye Research Institute, "much of what is published goes unchallenged, could be wrong, and the chances are that nobody knows."

Big brain: How much do fraud and error hurt science? Kohn concludes that fraud in important matters comes to light, while trivial fraud, which may not be caught, isn't damaging. Two cases undermine this argument. When fossil hunters in England unearthed Piltdown man—actually, just fragments of a large brain case and an apish jaw—they claimed it proved that a big brain was the first human feature to emerge. In fact, anatomical and chemical tests in 1953 revealed that the Piltdown fragments came from different creatures—the jaw was actually a modern orangutan's. In another scandal, British psychologist Cyril Burt reported data from twins starting in 1955 that indicated IQ is largely inherited. His results were cited as late as 1971 to argue that social standing is based on inherited intelligence. The subject is still hotly debated, but fresh scrutiny of Burt's papers in 1972 showed that he had at least "cooked" his twins data and at worst invented it outright.

Although the grant system and the pressure to publish may continue to invite fraud, there may be ways to minimize it. Braunwald himself suggests that labs adopt written procedures for conducting experiments and a system for investigating charges of fraud. Most scientists and journal editors agree. But, in fact, Stewart and Feder's paper remained unpublished for three years, as researchers whose work they criticized hired lawyers who, by calling the manuscript "defamatory," scared libel-conscious journals away from publishing it.

NEWSWEEK — AMERICAN MUSEUM OF NATURAL HISTORY

Deceptions: *Summerlin's dark mouse, Piltdown skull*

SHARON BEGLEY *with* MARY HAGER *in Washington and* SHAWN DOHERTY *in Boston*

APPENDIX 4

REPORTS ON THE DARSEE CASE FROM EMORY AND HARVARD UNIVERSITIES AND THE NATIONAL INSTITUTES OF HEALTH

I

REPORT OF *AD HOC* COMMITTEE TO EVALUATE RESEARCH OF DR JOHN R. DARSEE AT EMORY UNIVERSITY[1]

I. Brief Chronology of Events Leading to Investigation

At the annual meeting of the American Heart Association in November 1981 Dr Eugene Braunwald of Harvard Medical School informed Dr J. Willis Hurst and Dr Robert C. Schlant of Emory University that Dr John R. Darsee, a former resident and fellow in the department of medicine at Emory University had been accused of fabricating research data in the Cardiac Research Laboratory at Harvard. Dr Braunwald then wrote to Drs Hurst and Schlant on December 3, 1981 recounting the events at Harvard and urging an audit of Darsee's work at Emory. . . .

Dr Hurst promptly informed Dean James F. Glenn at Emory University School of Medicine by a letter dated December 11, 1981 of the events at Harvard Medical School. . . . A report of a committee composed of Harvard faculty and outside consultants, chaired by Dr Richard S. Ross, dean of the medical faculty and vice president for medicine, of the Johns Hopkins School of Medicine, was released on January 25, 1982, a copy of which was sent to Emory. . . .

On February 17, 1982 Dean Glenn appointed an *ad hoc* committee to review the work of Dr Darsee at Emory University. . . . The committee is composed of Dr Neil C. Moran, chairman of the department of pharmacology (chairman of the committee), Dr Elmer C. Hall, chairman of the department of biometry, Dr Charles M. Huguley, Jr, professor of medicine and Dr Panagiotis N. Symbas, professor of surgery.

In October 1981 the National Heart Lung and Blood Institute [NHLBI] staff became aware of problems involving the compatibility of data reported by investigators in a cooperative study of *Animal Models used in the Assessment of Therapy to Protect Ischemic Myocardium* (AMPIM). Some of the data submitted by the Harvard University group working under Dr Braunwald were different from those submitted by other centers and were not consonant with results obtained in several prior studies. The NHLBI staff learned from Dr Robert A. Kloner of Harvard that Dr Darsee had been found to fabricate data. Although Harvard was conducting its own study, the NHLBI decided to establish a small panel of senior and experienced cardiovascular investigators to inquire into the events at Harvard involving Dr Darsee and the AMPIM study. The panel completed its report on June 29, 1982 concluding that research data were fabricated in the course of work supported by NHBLI contract in the Cardiac Research Laboratory at Brigham and Women's Hospital, Harvard Medical School. The panel judged the fabrications to be the "acts of a single individual, Dr. John R.

[1] The addenda to this report include material and documents received by the investigative committee under assurances of confidentiality, and are therefore not reprinted here. [Editor.]

Darsee. . . ." The panel's report was released in February 1983 with a memorandum from the associate director of the NIH [National Institutes of Health] for extramural research and training to the director of NIH. The NIH memorandum made several recommendations. Among these are, one, that Dr Darsee be debarred from receiving financial assistance and contract funding from NIH for a period of ten years, two, that he be excluded from service on any NIH peer review or program advisory committee for a period of ten years, and three, "that NIH conduct a thorough review of Dr. Darsee's research and publications at Emory University." The last recommendation was based on information the NIH received after the panel had submitted its report that suggested that falsifications by Dr Darsee "were not isolated events but were part of a pattern involving Dr. Darsee's research at Emory University.". . .

II. Biographical Sketch of Dr John R. Darsee

Dr Darsee, a native of West Virginia, received the B.S. degree from the University of Notre Dame in 1970 and the M.D. degree from Indiana University School of Medicine in 1974. From 1974 to 1977 Dr Darsee was, in sequence, intern in medicine, junior resident in medicine and senior resident in medicine at the Emory University Hospitals. He was a clinical fellow in cardiology at the Emory University Hospitals from 1977 to 1979. In the year 1978–79 he was concurrently chief resident in medicine at Grady Memorial Hospital and instructor in medicine, Emory University School of Medicine.

In 1979 Dr Darsee went to Harvard Medical School as a research fellow in cardiology. In 1980 he was appointed instructor in the department of physiology and research associate in medicine at the Brigham and Women's Hospital. He had been offered a position in the department of medicine as assistant professor to be effective in July 1981, but the discovery of fabrication of data in the Cardiac Research Laboratory led Dr Braunwald to withdraw the offer of a faculty position and to termination of Dr Darsee's appointments as a fellow at Harvard Medical School and Brigham Hospital as well as his NIH postdoctoral fellowship.

Dr Darsee came to Emory University with strong recommendations from Indiana University School of Medicine. It was pointed out, however, that his performance in basic science courses was not distinguished but that he did do high quality work in his clinical years in medical school. . . . During the time Darsee was at Emory he was praised by nearly all the persons with whom he associated, whether faculty or medical students. His ratings by faculty and students during his internship and housestaff years were exemplary. . . . He was considered to be a brilliant person who had a prodigious capacity for work. We have noted no impressions of dishonesty among the persons who knew Dr Darsee at Emory except for Dr Ayten Someren of the department of pathology and laboratory medicine who noted mislabeled figures in a book chapter on cardiac neoplasia . . . and who noted discrepancies in Darsee sponsored abstracts for "the house staff research day" of the department of medicine . . . and also except for the recent statement of Dr Richard Stappenbeck in regard to the selection of

patients reported in the paper on mitral valve prolapse and ophthalmoplegia. . . .

Dr Darsee was highly recommended by Emory faculty to Dr Braunwald at Harvard. . . . Dr Braunwald and associates at Harvard also initially thought of Dr Darsee as "a prodigious worker who learned rapidly, was creative, well read, and excellent with his hands. Because of this extraordinary record and promising future he was accorded more and more responsibility and given the opportunity to work with less close supervision." (From page 2 of "Report of an *Ad Hoc* Advisory Committee to the Dean of the Harvard Medical School on Dishonesty in Scientific Research.")

Thus, until the recognition of falsification of data at Harvard, Darsee was considered to be a brilliant and outstanding physician, teacher and researcher. Certainly, prior to the disclosure at Harvard, none of the faculty at Emory, other than Dr Someren, appeared to recognize any trait or evidence suggesting dishonesty in his work.

III. Evaluation of Research of Dr Darsee at Emory University

A. Method of Procedure

The committee began with an evaluation of the papers of Darsee. A questionnaire was then sent to all of Darsee's collaborators at Emory in which information about the original records, writing and review of the papers and source of support for the research was sought. . . . After the questionnaires were returned the committee interviewed most of the collaborators. As a result of the interviews ramifications developed requiring interviews with other persons. Because there was no curriculum vitae and bibliography of Darsee at Emory and because we had no address at which to contact him, Dr Eugene Braunwald was called. He provided the curriculum vitae and bibliography and subsequently other information. Attempts to contact Dr Darsee by registered mail at a Boston address provided by Dr Braunwald were unsuccessful. When we learned of Dr Darsee having taken a position on the staff of Ellis Hospital in Schenectady, New York, we sent a registered letter to him there requesting the original records. He subsequently replied and has sent copies of records of some of the projects. . . . Dr Darsee's collaborators had provided information in the absence of original records. An interim report was submitted to the dean in December, 1982, at which time he instructed Darsee's collaborators at Emory to obtain expeditiously all information possible for this committee. . . .

B. Evaluation of Published Papers

Ten research papers were published as a result of the work of Dr Darsee at Emory University. The committee has found serious problems in seven of these papers. Each paper is discussed below in chronological order of publication. . . .

1. "Reversible severe congestive cardiomyopathy in three cases of hypophosphatemia." J. R. Darsee and D. O. Nutter. *Annals of Internal Medicine*, 89: 867–870, 1978.

The committee has not been able to see original data. The three patients described have not been identified because we do not have their names and because according to Dr Nutter "The Grady Hospital computer does not index medical records by diagnosis." However, recently Dr Robert Schlant, chief of cardiology at Grady Hospital, queried Dr Asa Yancey, medical director of the hospital, who stated that patients are indeed indexed in the computer by diagnosis.

Dr Nutter cites two papers, one an animal study and a human patient study, the results of which support the conclusions of this paper.

The project was not submitted to the Human Investigations Committee because it involved only review of charts. . . .

Recommendations

The editor of the *Annals of Internal Medicine* should be notified by letter from Dr Nutter, countersigned by Dr Glenn, that because the identity of the patients has not been discovered, because we do not have the original records (they are presumably in Dr Darsee's possession) and because of the evidence of widespread fraudulent practices identified in other papers, the work described in this paper should be considered to be of uncertain validity.

2. "Neurogenic skeletal myopathy in patients with primary cardiomyopathy." J. R. Darsee, D. O. Nutter, L. C. Hopkins and S. B. Heymsfield. *Circulation*, 59: 492–497, 1979.

The project described in this paper was initiated by Dr Nutter who "identified the patients and collected the clinical cardiovascular data." Dr Hopkins did the neurological examinations, the electromyograms and the muscle biopsies. Dr Heymsfield provided space in the Clinical Research Facility and reviewed the data. The work was nearly complete when Darsee was invited to join the project. Dr Nutter has discarded his "clinical records," but Drs Hopkins and Heymsfield state that they have original records. All collaborators attest to the validity of the data in this paper. In spite of the late entry of Darsee into the project he was given responsibility for writing this paper. All collaborators reviewed the manuscript prior to submission to the journal.

The project was supported in part by the NIH Clinical Research Facility Grant, RR–00039 and was approved by the Human Investigations Committee (IRB Project N–1126).

Recommendation

This paper appears to be valid. A letter might be sent to the editor of *Circulation*, countersigned by Dean Glenn, stating that this paper has been evaluated as part of the investigation of the work of Darsee, and the committee considers this work to be valid.

3. "Hypertrophic cardiomyopathy and human leukocyte antigen linkage: Differentiation of two forms of hypertrophic cardiomyopathy." J. R. Darsee, S. B. Heymsfield and D. O. Nutter. *New England Journal of Medicine*, 300: 877–882, 1979.

The project described in this paper is stated to have been initiated by Dr Darsee, who according to the collaborators identified many of the patients, collected and analyzed the clinical data and submitted blood samples to the tissue typing laboratory at Grady Hospital. Dr Nutter states that he identified some of the patients, reviewed clinical data on some patients and advised Dr Darsee and edited the manuscript Dr Darsee wrote. Dr Heymsfield states that his role was to help pay for the HLA [human leukocyte antigen] typing through the grant to the clinical research facility and to encourage Dr Darsee.

At the beginning of the committee's investigation neither of the collaborators had records of the patients and other persons described in the paper. Subsequently, the tissue typing laboratory at Grady Hospital provided a list of persons by name, age, sex, race and HLA types. However, the list gives no diagnoses. Another list, obtained as a NAPS Document deposited by Darsee with Microfiche Publications in New York, provides the age, sex, and HLA types of "70 probands with hypertrophic cardiomyopathy." No names are given, but the persons are placed in categories by race. The document from NAPS also gives pedigrees of five probands. Finally, we recently received from Dr Darsee copies of the reports of the tissue typing laboratory on which names are provided without diagnosis. . . .

Dr Nutter has searched the records of Grady Memorial Hospital and the Atlanta Veterans Administration Hospital and has analyzed the available data. He has found some minor and some major discrepancies. Among the major discrepancies are (1) the lack of laboratory records for the 70 matched normal control subjects described in the paper, (2) the diagnosis of hypertrophic cardiomyopathy cannot be confirmed from hospital records of at least six of the probands, (3) the correlations reported in the paper between laboratory and clinical data cannot be confirmed, and (4) laboratory data are not available for many of the members of the families described in the paper. In fact no laboratory data are available for the largest family the pedigree of which is shown in Figure 1 of the paper. The proband of this family does not appear in the NAPS listing of white probands. . . .

There are problems associated with the statistics. According to statements of both co-authors, the first version of the manuscript was returned by the editor of the journal with the suggestion of one reviewer that a statistical test called "lod scores" should be used. The committee was informed that Dr Darsee did not accept the suggestion that he seek advice from a member of the staff of the department of biometry. Instead, we were told, he read about and learned this statistical test and then applied it himself to the data. Dr Louis Elsas of the department of pediatrics and a geneticist states that Dr Nutter brought a table of data, similar to Table 3 of the paper, for evaluation. Elsas advised Nutter on the use of statistics but did not do the analysis himself. Although he is acknowledged in the paper for "reviewing the data," Dr Elsas states that he never met Darsee nor did he see the manuscript.

When the manuscript was resubmitted to the journal, presumably with

Darsee's calculations of lod scores and without the advice of a statistician, the paper was accepted.

There is no record that this project was submitted to the Human Investigations Committee for approval.

Conclusion

The evidence available provides strong indications that this paper contains data that are partially or completely fraudulent. The conclusion of the paper that "there is a heritable, nonhypertensive form of hypertrophic cardiomyopathy linked to the HLA loci on chromosome 6 and that a sporadic form is associated with severe, systemic hypertension" cannot be supported by the available evidence.

Recommendations

1. The editor of *The New England Journal of Medicine* should be notified by a letter from Drs Nutter and Heymsfield, countersigned by Dean Glenn, that the available evidence that has been evaluated by the committee fails to sustain the conclusions of the paper, and that the results of the entire paper must be considered to be invalid. The letter to the editor should provide a more detailed description of the problems, such as the analysis in the memorandum of Dr Nutter, dated March 16, 1983. . . .

2. The National Institutes of Health must be notified of the conclusion of this committee and informed that this clinical project was not submitted to the Human Investigations Committee for approval.

4. "Mitral valve prolapse and ophthalmoplegia: A progressive, cardioneurologic syndrome." J. R. Darsee, C. L. Miklozek, S. B. Heymsfield, L. C. Hopkins and N. K. Wenger. *Annals of Internal Medicine*, 92: 735–741, 1980.

This paper suggests that there is a "distinct neuromuscular disorder with metabolic and immunologic features [that is] associated with mitral valve prolapse and progressive mitral regurgitation." Dr Nanette Wenger suggested this study to Darsee and identified the initial two patients for the study. Of these two patients one is still a regular patient at the Grady Cardiac Clinics and the other died several years ago. Dr Wenger states the clinical data on the living patient are "completely correct," but she cannot remember the name of the second patient who died or that of a third patient from the Grady Clinics. . . . The original data were taken by Dr Darsee.

The paper describes seven cases. In the absence of names of patients it is not possible to verify diagnoses or laboratory data. Dr Heymsfield reports that he has reviewed the charts of two patients he recalled (but Dr Heymsfield does not give patient identification). However, he states that he cannot match these two patients with any of the seven reported in the paper, using as a guide the race, weight and height. Heymsfield also noted discrepancies in laboratory data in the patients' records and those reported in the paper. . . .

The paper reports that skin biopsies were done on six of the seven subjects

and on six matched control subjects. The paper further reports that biochemical and immunoglobulin studies were done on the cultured fibroblasts. The tissue culture studies were done at the tissue culture laboratory of the Georgia Mental Health Institute by Mr Martin Hollingsworth who no longer works there. He has reported by telephone that Darsee approached him requesting assistance in tissue culture. Hollingsworth cultured skin fibroblasts for Darsee but cannot remember whether the cells were brought by Darsee or were cells available in the laboratory. Hollingsworth gave Darsee flasks of cultured cells or pellets but does not know what Darsee did with these. Hollingsworth states that he did no more than culture the cells, that is, he did none of the biochemical or immunoglobulin studies. Hollingsworth worked for Dr John Galambos of the division of gastroenterology and was paid by Dr Galambos' NIH grant. . . . Dr Galambos reported by telephone that he did not know Hollingsworth was doing the work for Darsee and that neither Hollingsworth nor any other technician under Dr Galambos would have done other types of studies on the cells, such as the biochemical assays. The collaborators on this paper do not know who did the experiments on the cultured fibroblasts.

Dr Richard Stappenbeck who was a senior resident in neurology when Darsee was chief resident in medicine at Grady Hospital has charged in a letter . . . that "the data which he [Darsee] collected was spurious as relates to assigning these patients to the diagnostic classification of progressive external ophthalmoplegia." He states that in reviewing the clinic charts of "numerous of these cases" that "These patients manifested other problems such as stroke, effects following head trauma, and perinatal damage." Dr Stappenbeck states that he and other neurology residents did not think that Darsee should have included these patients in the group called progressive external ophthalmoplegia.

The study according to Dr Heymsfield was part of the project approved by the Human Investigations Committee on neurogenic skeletal myopathy and cardiomyopathy (IRB Project N-1126).

Conclusions

1. There is uncertainty about the identity of the patients in this report. No one at Emory has complete original records with patients' names.

2. Dr Heymsfield has found some discrepancies between the data in a few (unidentified) patients and the data in the paper.

3. Dr Stappenbeck's allegations are serious, but it is difficult to see how they can be proved in the absence of names of patients used in the study.

4. The issue of the fibroblast work suggests fabrication of data. That skin fibroblasts were cultured for Darsee seems clear on the basis of Martin Hollingsworth's testimony. However, at this time we do not know whether the cells were from samples of skin given to Hollingsworth by Darsee or were skin cells already under culture at the tissue culture laboratory. The person (or persons) who would have done the biochemical and surface immunoglobulin tests on the cells has not been identified. Therefore, we must question whether these tests were done at all.

Recommendations

1. A letter signed by Drs Heymsfield, Hopkins and Wenger and countersigned by Dean Glenn should be sent to the editor of the *Annals of Internal Medicine* stating that because the identity of the patients in this study could not be confirmed and because we have no evidence that the work on fibroblasts was done, the results and conclusions of the paper should be considered to be invalid. Dr Miklosek should be invited to sign the letter.

2. The NIH should be informed because the project was supported in part by the Clinical Research Facility grant and the tissue cultures were done under the auspices of Dr Galambos' grant. The NIH should also be informed that this study was not approved as such by the Human Investigations Committee, but appeared to have been done under the approval given for IRB Project No. 1126.

5. "Transcutaneous method of measuring Doppler cardiac output—I. Comparison of transcutaneous and juxta-aortic Doppler velocity signals with catheter and cuff electromagnetic flowmeter measurements in closed and open chest dogs." J. R. Darsee, J. R. Mikolich, P. F. Walter and R. C. Schlant. *American Journal of Cardiology*, 46: 607–612, 1980.

The original records of the experiments reported in this paper were taken to Boston by Darsee. Darsee has submitted a notebook of copies of tabular data from the experiments, but not the original tracings. . . . The collaborators have told the committee that Darsee made the Doppler flow tracings on a separate recording instrument independently from the tracings of the electromagnetic flowmeter tracings and from the thermodilution determinations of cardiac output. Darsee calculated cardiac output from the Doppler tracings and was given the data from the electromagnetic flowmeter tracings and the thermodilution data by the collaborators. Darsee then compiled the data and calculated correlation coefficients. The collaborators reviewed the compiled and analyzed data.

Dr Walter states that he was present during all of the dog experiments except for a period of two weeks while he was on vacation. During those two weeks Dr Mikolich worked in the laboratory with Darsee. . . . Dr Schlant has reviewed the notebook copies of data sent by Dr Darsee. . . . Dr Schlant has also provided the Committee with copies of correspondence to the *American Journal of Cardiology* dealing with this paper. . . .

Conclusions

1. The coefficients of correlation are high, suggesting that the data may have too good a fit. However, Dr Walter states another group has shown an R value of 0.97 comparing Doppler data and thermodilution derived stroke volume.

2. The absence of original records prohibits analysis of the data. The method by which Darsee was given electromagnetic flowmeter and thermodilution data to compare with his derived Doppler data would have given Darsee latitude for falsification of the data.

Recommendations

1. In the absence of original data and because of the evidence of widespread fraudulent practice on Darsee's part a letter should be sent to the editor of the *American Journal of Cardiology* signed by Drs Walter and Schlant, stating that the validity of the data in this paper cannot be verified. Dean Glenn should countersign the letter. Dr Mikolich should be invited to sign the letter.

2. The Georgia affiliate of the American Heart Association, a sponsor of this work, should be notified.

6. "Transcutaneous Doppler method of measuring cardiac output—II. Noninvasive measurement by transcutaneous Doppler aortic blood velocity integration and M mode echocardiography." J. R. Darsee, P. F. Walter and D. O. Nutter. *American Journal of Cardiology*, 26: 613–618, 1980.

The original records of the experiments reported in this paper were taken by Darsee. Darsee has submitted a notebook of copies of tabular data from the experiments, but we do not have the original tracings. . . . Dr Walter states that he was not present at all studies at the Veterans Administration Hospital and he is not sure how many of the patients were at the Veterans Administration Hospital and how many were at Grady Hospital. Dr Nutter supervised the experiments at Grady. Dr Darsee processed the Doppler data and compiled them with the data obtained from Fick, thermodilution and echocardiographic determinations. Walter and Nutter then reviewed the compiled data. They did not analyze the "raw" data from the Doppler tracings.

The question regarding the possibility that the coefficients of correlations are "too large" has been addressed by both Walter and Nutter who cite other studies that have high R values also.

Dr Schlant has reviewed the notebook of copies of data sent by Darsee. . . .

This project was approved by the Human Investigations Committee (IRB, 241–77).

Conclusions

The absence of original records precludes analysis of the data. The method by which Darsee was given data to compile and to compare with his own derived Doppler data flow would have given him latitude for falsification of the data.

Recommendations

1. In the absence of original data and because of the evidence of widespread fraudulent practice by Darsee, a letter should be sent by Drs Walter and Nutter to the editor of the *American Journal of Cardiology*, countersigned by Dean Glenn, stating that the validity of the data in this paper cannot be verified.

2. The Georgia affiliate of the American Heart Association, a sponsor of this work, should be notified.

7. "Decreased myocardial taurine levels and hypertaurinuria in a kindred with mitral valve prolapse and congestive cardiomyopathy." J. R. Darsee and S. B. Heymsfield. *New England Journal of Medicine*, 304: 129–135, 1981.

There are no original records of this project at Emory University. Dr Heymsfield states that he does not remember the subjects and thinks none were admitted to the Clinical Research Facility. Dr Heymsfield states that he "feels confident" that the Clinical Research Facility performed some of the blood analyses, but in the absence of names he cannot verify this. The notebook received from Dr Darsee has no names and no dates. The paper does not identify any hospital or clinic as the source of the patients.

The paper states that HLA types were obtained, but we have no record of these. The absence of original records and names precludes identification of the persons typed.

The paper states that myocardial tissue was obtained in three subjects four hours after death. Because no names of patients are available we cannot identify the possible persons. Furthermore, if the patients were at Grady Hospital it is highly unlikely that such "fresh" myocardial samples could have been obtained. Dr Ayten Someren, professor of pathology and director of anatomical pathology at Grady Memorial Hospital, stated emphatically that she would have known of the very improbable events of autopsies performed as early as within four hours of the patients' deaths.

The paper also states that myocardial samples were taken from the "anterior aspect of the left ventricle at the time of insertion of an epicardial pacemaker." There is no available source to verify that such biopsies were done at either Grady or Emory Hospitals. Such a procedure obviously requires approval of the Human Investigations Committee. There is no record of approval of this project. The paper also states that myocardial tissue samples "were also obtained at autopsy in the following control groups and processed in a similar fashion: 10 patients who had died with trauma but without heart disease (the 'normal' controls), who were matched for the age and sex of the five kindred subjects; and 12 unrelated patients who had died with idiopathic congestive cardiomyopathy." The implication here is that these hearts were also obtained within four hours of death. This would not have been possible. All of these myocardial samples were stated to have been obtained for microscopic analysis and for determination of amino acids, protein and hydroxyproline, the latter to estimate collagen. There is no record that any of these analyses were done.

The acknowledgement at the end of the paper expresses indebtedness "to Johnson S. Caulder, Ph.D., for performing the amino acid analysis, to Lawrence D. Bergman, M.D., for obtaining the heart tissue during pacemaker placement, to Myron C. Filstein, M.D., for referring several family members, and to Ms. Nancy Watterson for assistance with the manuscript."Ms Watterson is a typist in the Cardiac Research Laboratory at Brigham and Women's Hospital (personal communication from Dr Braunwald), but the other three persons are unknown at Emory University and at Harvard University. . . .

Dr Darsee was asked by mail about the identity of these persons. His reply

states "The names of the three persons acknowledged at the end of the 'Taurine' paper were mentioned in the hospital chart of one of the family members I first saw in clinic when I was an intern in 1974. They had contributed to the workup of this family prior to my involvement. Neither Steve Heymsfield nor I had any correspondence with them. They are not from Georgia but I don't remember where any of them are from.". . .

The paper states that samples of venous blood were obtained from all 43 members of the kindred "on two consecutive days after an overnight fast twice a year" and that twenty-four-hour urine samples were collected from "all family members . . . twice yearly or more frequently." It seems highly unlikely that these samples, especially the 24-hour urine samples, could be reliably collected from all 43 persons particularly from the two year old child.

In the absence of names of subjects to verify diagnoses the pedigree shown in Figure 1 is questionable.

We have no records of "taurine loading tests." Therefore, the data depicted in Figure 4 and Table 1 are questionable. Because there are no records of determination of myocardial taurine, hydroxyproline and collagen and because of the probability that the fresh myocardial tissue could not have been obtained, the data in Table 2 are questionable.

The paper is listed as coming from the department of medicine, Emory University School of Medicine and Harvard Medical School. However, there is no acknowledgement of collaborators from Harvard and no indication any of the study was done at Harvard.

There is no record of application to the Human Investigations Committee for approval of this project.

This work is stated to have been supported in part by NIH grant RR–00039.

Conclusions

1. There is no evidence that any of the work described in this paper is valid.

2. The use of fictitious names, the improbability, or impossibility, of Darsee having obtained myocardial tissue at autopsy within four hours of death in 25 subjects and the fact that fresh myocardial tissue said to have been obtained at pacemaker placement in two patients cannot be verified, the improbable aspect of getting reliable 24-hour urine samples on all subjects as stated, the lack of evidence of taurine-loading tests having been done at Emory or Grady Hospitals and the lack of application for approval of a clinical trial by the Human Investigations Committee all lead to the conclusion that the entire paper may be fictitious.

Recommendation

1. Dr Heymsfield should write a letter, countersigned by Dean Glenn, to the editor of *The New England Journal of Medicine* stating that the data in this paper are not valid and may be fictitious for the reasons stated above.

2. Notify the NIH of the lack of validity of this paper and the fact that there is no application on record for approval of the project by the Human Investigations Committee and that the work was supported by the NIH grant to the Clinical Research Facility.

8. "Hemodynamics of LeVeen shunt pulmonary edema." J. R. Darsee, J. T. Fulenwider, L. F. Rikkers, J. D. Ansley, B. F. Nordlinger, G. Ivey and S. B. Heymsfield. *Annals of Surgery*, 194: 189–192, 1981.

This project originated in the department of surgery. Dr Darsee came into the project late. His role according to Dr Heymsfield was primarily to write the manuscript and to participate in the study of one patient. . . . The paper was first sent to *Annals of Internal Medicine* and was rejected. Later, the paper was sent to *Annals of Surgery* and was accepted. Drs Heymsfield and Fulenwider (the only two authors still at Emory) attest to the validity of the paper.

The paper was supported by NIH grants RR–00039 (Clinical Research Facility) and AM–15736 (Dr Dean Warren, professor of surgery, principal investigator.)

The project was approved by the Human Investigations Committee (No. 220–77).

Conclusion

This paper appears to be valid.

Recommendation

No action is needed. A letter might be sent to editor of *Annals of Surgery*, countersigned by Dean Glenn, stating that the paper has been evaluated as part of the investigation of the work of Darsee, and the committee considers this paper to be valid.

9. "Paradoxical rise in left ventricular filling pressure in the dog during positive end-expiratory pressure ventilation. A reversed Bernheim effect." J. R. Darsee, J. R. Mikolich, P. F. Walter and R. C. Schlant. *Circulation Research*, 49: 1017–1028, 1981.

This paper is listed as coming from Emory University and Harvard University, but it does not acknowledge any of the faculty at Harvard. The collaborators at Emory state that the work was not completed at Emory and that Darsee told them he completed experiments at Harvard. Dr Braunwald has written to state that experiments on this project were not done at the Cardiac Research Laboratory at Harvard.

There are discrepancies in the descriptions in the paper and the reports of the collaborators. Dr Walter states that he was in the laboratory when the initial experiments were done. The atrial well (or reservoir) as described in the paper was not used in those experiments nor were ultrasonic crystals. Walter was told by Dr Mikolich that subsequently, when Walter was not in the laboratory, experiments were done with ultrasonic crystals to measure septal to free wall and anterior to posterior wall dimensions, but the atrial well was not used. Darsee then informed Walter and Mikolich that more experiments would be done at Harvard using another set of crystals that would measure base to apex dimension changes. Walter states that he reviewed the data and the manuscript and further states that "all the data presented in the manuscript that included the atrial well and apex to base dimensions would have been accumulated in Boston.". . .

Dr Mikolich, who is now in practice in Youngstown, Ohio, states that he participated in these experiments from April through June, 1979. He was also told by Darsee that additional experiments were done at Harvard. According to Mikolich the experiments at Emory utilized ultrasound crystals for two-dimensional measurement of ventricular size and that the pericardium was left open. The descriptions in the paper are of experiments in which the pericardium was left intact "as much as possible" and in which three dimensional measurements of ventricular size were reported. . . . Dr Mikolich has provided a copy of a letter from Dr Darsee to the associate editor of *Circulation Research* dated January 21, 1981 at the time a revised manuscript was submitted, stating that "we felt that the referee's suggestions were so beneficial that we decided to repeat the entire series of experiments in an additional sixteen dogs in order to provide the kind of meaningful data sought by the reviewers.". . .

Dr Schlant has reviewed the notebook of copies of data sent by Darsee. . . . Dr Schlant also has provided the committee with copies of correspondence dealing with this paper. . . . Dr Schlant also submitted copies of two drafts of the manuscript.

Conclusions

1. All of the experiments described in the paper must be fictitious inasmuch as they were not done at Emory University and were not done at Harvard University, or they represent a combination of real data from the Emory experiments combined with fabricated data. Whichever, the paper cannot be accepted as valid.

Recommendations

1. A letter stating that the data in the paper are invalid should be sent to the editor of *Circulation Research* signed by Drs Walter and Schlant and countersigned by Dean Glenn. Dr Mikolich should be invited to sign the letter also.
2. The Georgia affiliate of the American Heart Association should be notified of the decision.

10. "Prevalence of mitral valve prolapse in presumably healthy young men." J. R. Darsee, J. R. Mikolich, N. B. Nicoloff and L. E. Lesser. *Circulation*, 59: 619–622, 1979.

None of the authors of this paper are now at Emory so that we do not have access to records of the data. Dr Walter states that although he was not involved in the project, he was at the Veterans Administration Hospital when the study was done. Walter states that Dr Lesser, who supervised the project and was on the faculty at the time of the study, assured Dr Walter that he looked at every echocardiogram and phonocardiogram. Walter states that he believes there is no problem with this paper. . . .

Conclusion

In the absence of data and names of the subjects, it is not possible to verify the conclusions of this paper. However, because of Dr Walter's statement we probably can conclude that the data are valid.

Recommendation

Contact Dr Lesser to see if he has original research data to allow a critical evaluation. If neither he nor the other co-authors have the data, the editor of the journal should be notified that the data may not be valid.

C. EVALUATION OF BOOK CHAPTERS

1. "The electrocardiogram." J. R. Darsee and J. W. Hurst, in *Clinical Methods: The History, Physical and Laboratory Examination*. 2nd Ed. Walker, H. K., Hall, W. D. and Hurst, J. W. (Editors) (Boston: Butterworths, 1980), pp 1122–1161.

Dr Hurst states that Darsee designed Figure 211, many of the other figures being from a previously published book. In Figure 211, the mean T vector is drawn as being inferior to a perpendicular lead AVF whereas Dr Hurst thinks it should be drawn more superiorly. . . .

Conclusion

The location of the vector in one illustration is a minor problem that will be corrected in the next edition of this book.

Recommendation

No action need be taken.

2. "Hypophosphatemic myocardial dysfunction: The role of phosphorus and high-energy phosphate compounds in myocardial function and dysfunction." J. R. Darsee in *The Heart*. Update III. pp.1–9. Hurst J. W. (Editor) (New York: McGraw Hill, 1980).

This is a review of the role of phosphate compounds in myocardial function. We have no statement from Dr Hurst about this chapter. The committee has not critically reviewed it. It should be noted that Darsee cites his own work on hypophosphatemia (see evaluation of paper in *Annals of Internal Medicine* in Section III, B, p. 4, and of abstract in Section III, D, 2, p. 17), including two figures from his paper on this subject. Because the paper in *Annals of Internal Medicine* is of questionable validity, the part of this chapter should be considered to be invalid.

Recommendation

This topic should be revised by another author for the next edition of this book.

3. "The manifestations of neoplasia on the heart, including nonmetastatic effects and the cardiotoxic effects of antineoplastic therapy." J. R. Darsee, J. W. Hurst and N. K. Wenger, in *The Heart*. Update IV. pp 219–86. Hurst, J. W. (Editor) (New York: McGraw Hill, 1981).

This book contains three pictures, the legends of which are in error. These were first recognized by Ayten Someren, director of anatomical pathology at Grady Hospital. Figure 11, p. 236, is described as an "intra-cavitary liposarcoma in situ in a patient who died suddenly with no previous

symptoms and signs." This figure, taken from the Grady Hospital files (A79–066), actually shows a metastatic fibrous histiocytoma in the heart of a patient who had a primary soft tissue fibrous histiocytoma removed five years previously and who died of recurrent tumor. Figure 14, p. 238, is labeled as "carcinoma of the breast which had metastasized to epicardial fat of the right ventricle." This figure (A73–238 of the Grady files) is actually of a small calcified epicardial lipoma found incidentally at autopsy. Figure 23, p. 245 is labeled "metastatic adenocarcinoma (unknown primary) involving ventricular endocardium, growing into the ventricular cavity and out the pulmonary artery (arrow)". The picture (A79–013 from Grady files) actually shows a large, partly necrotic bronchogenic carcinoma involving the upper and middle lobes of the right lung, which is extending into and obstructing the superior vena cava. It does not show the pulmonary artery or an open right ventricle as the legend describes. . . .

Dr Nanette Wenger states that these are the only errors that she and Dr Hurst can identify in the chapter.

Conclusion

There are obviously wrong descriptions of pictures of autopsy material from the Grady Hospital files. These figures were used by Darsee and credit is given to Mr Eugene Semples, an associate who works in the autopsy laboratory, instead of to Dr Someren. She did not review these slides with Dr Darsee. It is not possible to tell whether these errors in the labels are accidental or by design.

Recommendation

Correct in next chapter of the book.

4. *Other book chapters listed in Dr Darsee's bibliography*

There are eleven chapters listed in the bibliography that were never published. These were to be part of a *Manual of Cardiology for House Officers* to be edited by Drs Hurst, Darsee and Miklozek. The book was ready for publication when Dr Hurst learned of the problems at Harvard. Dr Hurst promptly told the publisher that the book could not be published. . . .

D. EVALUATION OF ABSTRACTS

1. "Skeletal myopathy in patients with primary cardiomyopathies." J. R. Darsee, D. O. Nutter, L. C. Hopkins and S. B. Heymsfield. *Circulation*, 56: III–218, 1977.

This abstract is related to paper on neurogenic skeletal myopathy as described in Section III, B, 2. . . . This project was approved by the Human Investigations Committee (No. N–1126).

Conclusion and Recommendation

Data are apparently valid. No action is recommended.

2. "Noninvasive analysis of PEEP-induced hemodynamic changes by Doppler stroke volume ratios." J. R. Darsee and Donald O. Nutter. *Chest*, 74: 334, 1978.

According to Dr Nutter, Darsee initiated this project and collected and analyzed the Doppler flow data. Nutter reviewed the data. Nutter has no original data. There is no record of approval of this project by the Human Investigations Committee. Nutter states that Darsee may have assumed that approval of project No. 241–77 related to this project also.

Conclusion and Recommendation

In absence of original data and names of patients verification of data is impossible. Letter should be sent to the editor of *Chest* indicating that because of evidence of widespread fraudulence by Darsee and because of absence of original records we must assume the data may not be valid.

3. "Avidity of myocardium for phosphorus and calcium in hypophosphatemia." J. R. Darsee, T. G. Murray, S. B. Heymsfield and D. O. Nutter. *Clinical Research*, 26: 773A, 1978.

Both Dr Nutter and Dr Heymsfield state that this was a study done by Darsee and a medical student (Murray) under Heymsfield's direction. The work is related to the clinical report on hypophosphatemic cardiomyopathy. (See Section III, B, 1). This abstract is cited in the chapter on hypophosphatemic myocardial dysfunction in the book *The Heart*. . . .

Conclusion and Recommendation

In the absence of original records it is not possible to verify the validity of the data and the conclusions. Drs Nutter and Heymsfield should send a letter to the editor of the journal indicating that the data may not be valid.

4. "Mitral valve prolapse progressing to congestive cardiomyopathy in a large kindred with hypertaurinuria." J. R. Darsee and S. B. Heymsfield. (Intro. by J. W. Hurst). *Clinical Research*, 26: 762A, 1978 and *Clinical Research*, 27: 274A, 1979.

This abstract is related to the paper on taurine and the heart in *The New England Journal of Medicine*. See review of that paper in Section III, B, 7 . . . of this report. Note that this abstract was published twice in different years.

Conclusion and Recommendation

The data in this abstract cannot be considered to be valid. Dr Heymsfield should write to the editor of the journal explaining why.

5. "Hemodynamics of LeVeen-valve pulmonary edema." J. R. Darsee, J. D. Ansley, J. T. Fullenweider, L. F. Rikkers, B. Nordlinger and S. B. Heymsfield. *Clinical Research*, 26: 759A, 1978 and *Clinical Research*, 27: 264A, 1979.

This abstract is related to the paper on this subject in the *Annals of Surgery*. See Section III, B, 8, . . . for review of this paper. Note that this abstract was published twice in different years.

Conclusion and recommendation

Data are probably valid. A letter might be sent to the editor of *Clinical Research*, countersigned by Dean Glenn, stating that this abstract has been evaluated as part of the investigation of the work of Darsee, and the committee considers this abstract to be valid.

6. "Noninvasive quantification of mitral regurgitation by echo and Doppler stroker volume measurement." J. R. Darsee, D. A. Bray and D. O. Nutter. *Circulation*, 58: II-42, 1978.

Dr Nutter states that Darsee initiated this project and collected the Doppler and echocardiogram data in relationship to cardiac catheterization studies at Grady Hospital. Bray was a medical student participating in the study. Dr Nutter has no records from this study. The study itself did not receive Human Investigations Committee approval, but according to Nutter, Darsee may have assumed that he had approval through project No. 241-77. It is related to the paper discussed in Section III, B, 6, . . .

Conclusion and Recommendation

In the absence of names of patients and records of the experiments one must question the validity of the data. Dr Nutter should write to the editor of the journal explaining why the data may not be valid.

7. "Reversible, severe congestive cardiomyopathy from hypophosphatemia." J. R. Darsee and D. O. Nutter. *Circulation*, 58: II-23, 1978.

This abstract is related to the paper by the same name described in Section III, B, 1, ... Approval of Human Investigations Committee was not requested because it was a "chart" study that does not require approval.

Conclusion and Recommendation

Because the names and identities of the patients are not known, the validity of this study is questionable. The editor of the journal should be so informed.

8. "Human leukocyte antigens and hypertrophic cardiomyopathy." J. R. Darsee, D. O. Nutter and S. B. Heymsfield. *Clinical Research*, 26: 34A, 1978.

This abstract relates to the paper on hypertrophic cardiomyopathy and HLA linkage discussed in Section III, B, 3. There was no approval of this project by the Human Investigations Committee.

Conclusion and Recommendation

Because of the serious questions about the validity of the data as discussed for the full paper, the editor of the journal should be notified that the data in this abstract are probably not valid.

9. "Cardiac output by transcutaneous Doppler cardiosonography in myocardial infarction." J. R. Darsee and D. O. Nutter. *Clinical Research*, 26: 226A, 1978.

This abstract is related to the paper on Doppler flow measurement in patients described in Section III, B, 6, ... The project had Human Investigations Committee approval (No. 241–77).

Conclusion and Recommendation

Because of the problems discussed in relationship to the full paper the editor of this journal should be informed that the data in this abstract are of uncertain validity.

10. "Complement-resistant lymphocytes in mitral valve prolapse." J. R. Darsee, C. L. Miklozek, S. B. Heymsfield and D. O. Nutter. *American Journal of Cardiology*, 43: 367, 1979.

Both Dr Nutter and Dr Heymsfield state that they knew nothing about this project or abstract.

Conclusion and Recommendation

Because neither of the co-authors now at Emory know about this project and do not have original records, we must assume that the data are not valid. There is no record of Human Investigations Committee approval of the project, although Dr Nutter states it may have had approval under No. 311–78, "Autoimmune aspects of the mitral valve prolapse syndrome." A letter should be sent to the editor of the journal by Drs Nutter and Heymsfield stating that they have no knowledge of this project and that their names were put on the abstract without their permission or knowledge. Dr Miklozek should be invited to sign the letter.

11. "Noninvasive assessment of relative inotropic state by analysis of Doppler velocity curves." J. R. Darsee, P. F. Walter and D. O. Nutter. *Clinical Research*, 27: 161A, 1979. Also in *Circulation*, 60: II–124, 1979.

Dr Walter states that he played no role in this project nor did he see the abstract. Dr Nutter states that the study was performed and reported while he was on leave of absence from Emory. Darsee sent results to Nutter to review and Nutter advised on interpretation of data. Dr Braunwald comments on this abstract "I find this . . . shocking because 15 patients with 'overt cardiac failure' were given intravenous colloid while a catheter was placed in the pulmonary capillary position.". . . There is no record of application to the Human Investigations Committee for approval of this project.

Conclusion and Recommendation

Drs Nutter and Walter should send a letter to the editor of the journal stating that they either had no role in this project and knew nothing about it (Walter) or that only advice was given by mail and there was no participation in the study (Nutter), and that there was no Human Investigations

Committee approval. Because of these factors the data cannot be considered to be valid.

12. "Improved myocardial performance with oral levodopa in congestive cardiomyopathy." J. R. Darsee, S. B. Heymsfield, D. O. Nutter. *Clinical Research*, 27: 230A, 1979.

Dr Heymsfield states he knew nothing about the project or the abstract. Dr Nutter states that the study was performed while he was on leave of absence. Nutter was informed of the results and gave advice, but he played no direct role in the study. There is no evidence of application for approval of the project by the Human Investigations Committee.

Conclusion and Recommendation

Because the co-authors were not directly involved in the study, because the records are not available and because the project was not approved by the Human Investigations Committee, Drs Nutter and Heymsfield should send a letter to the editor of the journal stating that the results may not be valid.

13. "Cirrhotic cardiomyopathy: The hyperdynamic-unloaded-failing heart." J. R. Darsee, S. B. Heymsfield and D. O. Nutter. *Clinical Research*, 27: 265A, 1979. Also in *Circulation*, 60: II–38, 1979.

Dr Nutter played no direct role in this study which was conducted while he was on leave of absence from Emory. Nutter states that he advised Darsee on experimental design. He did not review the data or the abstract. Heymsfield states "Our facility had data similar to those appearing in the abstract on file in about 50 patients. I sent them to Darsee in the hope he would find them interesting. He also had cases of his own, which he added to complete the report. I know nothing more about the study." There is no evidence that this project received approval by the Human Investigations Committee.

Conclusion and Recommendation

In the absence of original records and of approval of the project by the Human Investigations Committee, one must conclude that the data may not be valid. A letter so stating should be sent to the editor of the journal by Drs Nutter and Heymsfield.

14. "Instantaneous length-tension-velocity relations: Assessment of left ventricular performance with a transcutaneous Doppler velocity flowmeter." J. R. Darsee, P. F. Walter and D. O. Nutter. *Circulation*, 60: II–41, 1979.

Dr Walter states he had no role in this project and learned of it only after the abstract was published. Dr Nutter states that he played no part in the project which was conducted while he was on leave of absence. Darsee informed Nutter of the final results and the abstract. There is no record of approval of the project by the Human Investigations Committee although

Dr Nutter states Darsee may have considered it to be under the approval of the approved project No. 241–77.

Conclusion and Recommendation

There is no evidence that the data in this abstract are valid. Drs Nutter and Walter should write to the editor of *Circulation* stating this.

15. "Effects of calcium and beta-blockade on the negative inotropic response to respiratory and metabolic acidosis in the normally contracting human heart." J. R. Darsee and D. O. Nutter. *Circulation*, 60: II–90, 1979.

Dr Nutter states that he did not participate in this project and did not see the abstract until it was published. There is no evidence of approval of the Human Investigations Committee. . . .

Conclusion and Recommendation

There is no evidence that the data described in this abstract are valid. Dr Nutter should write to the editor of the journal explaining why they are not valid.

16. "Centripedal hypertrophy: Explanation for outflow tract obstruction in hypertrophied hearts: Morphometric, cytometric, and chemical analysis." J. R. Darsee and D. O. Nutter. *Circulation*, 60: II–149, 1979.

Dr Nutter did not participate in this project and saw the abstract only after it was published. There is no evidence of Human Investigations Committee approval. It is highly unlikely that the studies on such a large number of hearts were done at all. There is no evidence that the biochemical studies reported herein were done at Emory University or Grady Memorial Hospital.

Conclusion and Recommendation

The absence of original records and the fact that Dr Nutter did not know of this project and the improbability of so many determinations having been done on relatively fresh hearts suggests that these data are not valid and are probably fictitious. Dr Nutter should write to the editor of the journal and indicate why the data are invalid.

17. "The heart in Crouzon's disease: A clinical, hemodynamic and histopathological description." J. R. Darsee and T. L. Feng. *Circulation*, 59/60 (Suppl. 2): 984, 1979.

The identity of T. L. Feng is not known. This abstract is not credible. . . .

Conclusion and Recommendation

It is unlikely that this project was done. The abstract is probably a fabrication. The editor of the journal should be notified.

18. "Cardiac output by transcutaneous Doppler blood flow measurement." J. R. Darsee, P. F. Walter and D. O. Nutter. *Circulation*, 58: II–174, 1978.

This abstract is related to the full papers discussed in Sections III, B, 5 and 6.

Conclusion and Recommendation
Notify editor of uncertain validity of data.

19. "Absence of myocardial fibroblast surface antigen in idiopathic congestive cardiomyopathy: Evidence for a viral etiology and importance of afterload." J. R. Darsee, S. B. Heymsfield and D. O. Nutter. *Chest*, 76: 371, 1979.

Dr Heymsfield states that he was unaware of this abstract until Dr Hurst brought it to his attention. He had no knowledge of this project. We have no statement from Dr Nutter.

Conclusion and Recommendation
In the absence of data and names of patients, the validity of the data in this abstract must be questioned. The editor of the journal should be notified.

20. "Myocardial ATP depletion and reversible cardiac dysfunction in the chronically hypophosphatemic conscious dog." J. R. Darsee, P. F. Walter and D. O. Nutter. *Circulation*, 62: III–177, 1980.

Neither Dr Walter nor Dr Nutter were part of this study. Dr Walter became aware of it only after publication of the abstract. Nutter states that Darsee informed him of the results of the work and gave the impression that part or all of the work was performed at Harvard. There is no evidence that the experiments described in this abstract were done. Darsee had no known access to a laboratory either to have the ATP analyses done or to do them himself. Dr Braunwald . . . . states that he and his group had developed a technique for myocardial biopsy of the type described in this abstract. When Darsee arrived at the Harvard laboratory, he did not know how to do this procedure. Therefore, he could not have done these experiments at Emory.

Conclusion and Recommendation
There is no evidence that these experiments were done either at Emory or at Harvard. Therefore, it must be concluded that they are probably fictitious. Dr Walter and Dr Nutter should write to the editor of the journal to explain why this abstract is invalid.

21. "Two distinct catecholamine cardiomyopathies from norepinephrine and epinephrine secreting pheochromocytomas." J. R. Darsee and D. O. Nutter. *Circulation*, 62: III-317, 1980.

Dr Nutter states that this study was performed while he was on leave of absence. Darsee sent Nutter parts of the data for comment and interpretation, but Nutter was not aware of the abstract until it was published.

Conclusion and Recommendation

There is no evidence that this study was done at Emory. The data are probably not valid. The editor of the journal should be so informed by Dr Nutter.

22. "Evidence for cardiac dysfunction following hepatitis B infection." J. R. Darsee, P. F. Walter, C. L. Miklozek, C. B. Bell and S. B. Heymsfield. *Circulation*, 62: III–318, 1980.

Drs Walter, Bell and Heymsfield knew nothing about this project or the abstract until it was published. The absence of evidence that the study was done suggests that the data are fictitious.

Recommendation

Drs Walter, Heymsfield and Bell should write to the editor stating the data are probably fabrications. Dr Miklozek should be invited to sign the letter.

23. "Relationship between response to inotropic agents and homeometric autoregulation of the myopathic heart." J. R. Darsee, C. L. Miklozek, S. B. Heymsfield and D. O. Nutter. *Clinical Research*, 28: 163A, 1980.

Dr Heymsfield states that he was completely unaware of this abstract until informed of it by Dr Hurst. He did not participate in this study. Dr Braunwald questions whether a Human Subjects Committee "would allow patients with congestive cardiomyopathy to receive no medication and then undergo general anesthesia for research purposes.". . .

Conclusion and Recommendation

There is no evidence that this study was done. It may be fabricated. Drs Nutter and Heymsfield should write to the editor of the journal, inviting Dr Miklozek to join them.

24 "Dissimilar cardiovascular effects of norepinephrine and epinephrine secreting pheochromocytomas: Documentation of two distinct catecholamine cardiomyopathies." J. R. Darsee and D. O. Nutter. *Clinical Research*, 28 (No. 2): 257A, 1980.

No statement has been received from Dr Nutter. . . .

Conclusion and Recommendation

It is unlikely that these studies were done. The editor of the journal should be notified by Dr Nutter of uncertain validity of the data.

25. "Potential detrimental effects of continuous positive pressure ventilation on hearts with subcritical coronary stenosis." J. R. Darsee, P. E. Walter, J. R. Mikolich and R. C. Schlant. *Clinical Research*, 28: 653A, 1980.

There is no statement from Dr Walter or Dr Schlant. There is no evidence that this study was done at Emory. . . .

Conclusion and Recommendation

It is doubtful that the data reported in this abstract are valid. The editor of the journal should be notified.

26. "The heart in pancreatitis: A combined clinicopathologic study of 128 cases." J. R. Darsee, R. F. Fincher, M. Fincher, C. L. Miklozek, A. O. Someren and J. W. Hurst. *American Journal of Cardiology*, 45: 395, 1980.

Dr Someren states that she was completely unaware of this abstract. She was not consulted. It is related to an abstract of a paper given before the Department of Medicine house staff research day in 1979. See discussion of the latter in Section III,E,1, . . .

Conclusion and Recommendation

The data in this abstract appear to be a gross elaboration of no more than 20 patients Dr Someren discussed with Dr M. Fincher for the research day. The editor of the journal should be notified that the data in this abstract are of questionable validity.

27. "Is nonfamilial asymmetric septal hypertrophy a different histological disorder?" J. R. Darsee, S. Anton, C. L. Miklozek and A. O. Someren. *American Journal of Cardiology*, 45: 492, 1980.

This abstract is related to the abstract of a paper given at the department of medicine research day in 1979 by Dr Anton. See discussion of latter in Section III,E,2, . . . Dr Someren was unaware of the existence of this abstract and did not participate in the project.

Conclusion and Recommendation

The data in this abstract cannot be verified. The editors of the journal should be notified.

28. "Normalization of pyruvate decarboxylation in cultured fibroblasts of patients with ophthalmoplegia-prolapse using a pyruvate dehydrogenase complex enzyme." J. R. Darsee, C. L. Miklozek, L. C. Hopkins, Jr., S. B. Heymsfield, N. K. Wenger and D. O. Nutter. *Clinical Research*, 29: 428A, 1981.

Drs Hopkins, Nutter, Heymsfield and Wenger state that they knew nothing about this abstract. The title of the abstract relates it to the paper entitled "Mitral valve prolapse with ophthalmoplegia: A progressive cardioneurologic syndrome," published in the *Annals of Internal Medicine*. See Section III,B,4, . . . for analysis of that paper and the issue of the tissue culture studies.

Conclusion and Recommendation

Because none of the co-authors who are now at Emory knew of this abstract, because we cannot identify the patients, and because there is doubt that biochemical analyses of the tissue cultures were done, the data in this abstract must be considered invalid. Drs Hopkins, Heymsfield, Wenger and

Nutter, joined by Dr Miklozek if she agrees, should write a letter to the editor of the journal explaining why the data in this abstract cannot be considered to be valid.

29. "Improvement in right ventricular function and cardiac output following right coronary artery occlusion with intrapulmonary artery balloon counterpulsation." J. R. Darsee and P. F. Walter. *Circulation*, 64: IV–68, 1981.

Dr Walter states that he and Darsee had discussed the desirability of having some right ventricular assist device. However, he did not take part in any such study and was not aware of such a study being conducted. His first knowledge of the work was through the publication of the abstract. There is no evidence that Darsee conducted these experiments at Emory University.

Conclusion and Recommendation

In the absence of evidence that the experiments were done it must be concluded that the data are fictitious. Dr Walter should write to the editor of the journal to explain why this abstract is invalid.

30. "Time course and severity of mitral regurgitation from ischemic papillary muscle dysfunction: Effects of balanced alpha and beta adrenergic blockade." J. R. Darsee, C. L. Miklozek and D. O Nutter. *Circulation*, 64: IV–210, 1981.

Dr Nutter states that he knew nothing about this project or the abstract until its publication. There is no evidence that Darsee performed these experiments at Emory. Nor is there evidence that he knew and used a technique for measuring papillary muscle length (and flow) *in situ*.

Conclusion and Recommendation

The experiments described here were not done at Emory University or Harvard University. Therefore, one must conclude that the abstract is probably fabricated. Dr Nutter should write to the editor of the journal, inviting Dr Miklozek to join him, to explain why this abstract is not valid.

31. "Endocardial fibroelastosis, coronary endotheliosis, and mitral valve hooding in Sandhoff's Disease: Evidence for uncontrolled protein production by cardiac fibroblasts." J. R. Darsee and S. B. Heymsfield. *Circulation*, 64: IV–280, 1981.

Dr Heymsfield states that he knows nothing about this project or the abstract. There is no evidence that this study was done at Emory or elsewhere. The descriptions of obtaining myocardial and coronary artery samples "within 2 hours after death" from four children with Sandhoff's disease and from children dying from other causes is highly unlikely and would not have been done at Emory. Furthermore, it is unlikely that he studied incorporation of radiolabeled lysine into fibroblasts grown from these myocardial and artery samples.

Conclusion and Recommendation

The probability that this project was done at all is extremely slim. It must be concluded that the abstract is probably entirely fictitious. Dr Heymsfield should write to the editor of the journal to explain this.

32. "Replacement of myocardial norepinephrine stores and improvement in cardiac performance in chronic heart failure by administration of levodopa." J. R. Darsee, J. R. Mikolich and D. O. Nutter. *Circulation*, 64: IV–69, 1981.

Dr Nutter states that he knew nothing about this project or abstract. There is no evidence that these experiments were done at Emory University.

Conclusion and Recommendation

Inform the editor of the journal of the likelihood that the data are not valid.

33. "Right ventricular ischemia and abnormal contraction: Potential deleterious effects of continuous positive pressure ventilation with minimal coronary narrowing." J. R. Darsee, P. F. Walter, J. R. Mikolich and R. C. Schlant. *Clinical Research*, 29: 444A, 1981.

This abstract was discovered as the report was being completed. Drs Walter and Schlant could not be reached to enquire about their knowledge of it. It is probable that this abstract is fictitious. It is dated 1981, is listed as from Harvard Medical School with no mention of Emory University and describes experiments on 18 dogs that involve multiple complex instrumentation. The work was not done at Emory, and it is not likely that it was done at Harvard.

Conclusion and Recommendation

This abstract is most likely a fabrication. The editor of the journal should be so notified.

E. EVALUATION OF ABSTRACTS SPONSORED BY DR DARSEE FOR DEPARTMENT OF MEDICINE HOUSE STAFF RESEARCH DAY

Two abstracts of papers presented at the annual department of medicine house staff research day are relevant to the problem of Dr Darsee's approach to research at Emory. The issues discussed below were brought forward by Dr Ayten Someren of the department of pathology and laboratory medicine.

1. "Armored hearts and non-coronary infarctions: unexplained features of pancreatitis." M. E. Fincher (resident in medicine). Sponsors: J. R. Darsee, A. O. Someren. During the 1978–79 year Dr Michael E. Fincher, a resident in medicine, requested the help of Dr Someren in a review of cardiac histopathology in cases of pancreatitis. She recalls reviewing and discussing with Dr Fincher the slides of approximately 20 cases of pancreatitis from the files of Grady Hospital. A few days before research

day she was shown an abstract prepared by Drs Fincher and Darsee which was included in the collection of abstracts compiled for the day. She discovered discrepancies between the findings she and Dr Fincher had reviewed and the findings described in the abstract. Furthermore, the abstract stated that "86 cases of pancreatitis were available for review and further study from the necropsy files since 1965." Also the abstract described "armored hearts" which according to Dr Someren is not a valid description based on her review of the 20 cases. Although she had agreed to serve as sponsor for Dr Fincher, she was surprised that she had not been consulted about the final abstract as well as the increase in number of cases and a different interpretation from hers. She informed Dr Hurst of the problem. He wrote a letter to Dr Fincher, but Dr Someren heard nothing more about the abstract until this committee brought to her attention the mislabeled figures in the chapter in the Update IV issue of *The Heart* (see Section III,C,3.) She then told us of the problem with this abstract and the one discussed below. Later, an abstract was discovered in *The American Journal of Cardiology* that appeared to be a "compilation of two research day abstracts" presented by Dr Fincher. *The American Journal of Cardiology* abstract describes 128 patients. Dr Someren reviewed only 20 patients with Dr Fincher. (See discussion of abstract No. 26 in Section III,D,26, . . .

2. "Absence of septal myofiber disarray in hypertensive hearts with asymmetric septal hypertrophy." S. L. Anton (resident in medicine). Sponsors: J. R. Darsee, A. O. Someren. This abstract was not reviewed by Dr Someren nor did Dr Anton consult with her about the cases although he had asked her if she would sponsor the work. She learned of the abstract only a few days before the research day. She informed Dr Hurst, who wrote a letter to Dr Anton. Subsequently, she learned that the same material was presented in slightly modified form by Dr Darsee in an abstract in *The American Journal of Cardiology*. The modification was an increase in number of cases from 78 in the Anton abstract to 91 cases in the Darsee abstract. (See discussion of abstract No. 27 in Section III,D,27, . . .)

IV. Discussion

From the time Dr Darsee was a third year medical student at Indiana University School of Medicine until May 1981 when he was found to be fabricating data in the Cardiac Research Laboratory at Harvard Medical School he was generally considered to be a bright, energetic, honest young man with a prospect of a brilliant future in academic medicine. Certainly few people at Emory suspected that this brilliant young man was capable of dishonesty in his research activities. Most persons were impressed by his prodigious energy and capacity for work.

A pattern of deception, manipulation of data and of people and outright fabrication of research data now appears to be a truer characterization of this person. Of nine full papers published while at Emory only two can be considered to be valid. One of the papers, that on taurine and the heart

(Section III,B,7, . . .) is the most flagrant example of the fabrication and dissembling that pervades most of the papers and abstracts. A structure was built that would appear to be convincing evidence for the conclusion as written in the paper "hypertaurinuria and depressed levels of taurine in the myocardium may be associated with one type of mitral-valve-prolapse syndrome and a rapidly progressive form of congestive cardiomyopathy." The kindred of 46 members he describes cannot be identified, and therefore, the pedigree shown in Figure 51 of the paper is meaningless. The claim to have obtained hearts within four hours after death for amino acid analyses has no substance. The taurine loading tests were not done at Emory University, and therefore, were probably not done at all. The persons acknowledged for performing amino acid analyses, for providing fresh myocardium from two patients at pacemaker placement and for referring several family members cannot be identified at Emory University, Harvard University, Atlanta or Boston. Dr Darsee in response to a query about these persons replied that he found their names "in the hospital chart of one of the family members I first saw in clinic when I was an intern in 1974. . . . They are not from Georgia but I don't remember where any of them are from." Two tables have detailed data representing means and standard deviations of serum and urinary taurine, creatinine and taurine clearances and myocardial taurine, hydroxyproline and collagen. In the absence of evidence that these analyses were done under auspices of Emory University and in view of the fact that the amino acid analyses are attributed to the unidentified Johnson S. Caulder, Ph.D., we are forced to conclude that these data are fabricated.

The paper on cardiomyopathy and human leukocyte antigen types (Section III,B,3, . . .) is a less flagrant but no less serious example of the dissembling and fabrication. This paper concludes that "there is a heritable, nonhypertensive form of hypertrophic cardiomyopathy linked to the HLA loci on chromosome 6 and that a sporadic form is associated with severe, systemic hypertension.". . . we cannot find evidence to support this contention. That HLA typing was done on a large number of people is clear, but the pedigrees of the families and the alleged linkage of hypertrophic cardiomyopathy with specific human leukocyte antigens cannot be confirmed. One must conclude, therefore, that the results are not valid and that the paper is fraudulent.

The paper on paradoxical rise in left ventricular filling pressure during positive end-expiratory pressure ventilation (Section III,B,9) exemplifies the dissemblance and fabrication that characterize Darsee's activities. Experiments were performed at Emory on dogs in collaboration with colleagues, but perhaps because the results were not as expected or because more work was needed, Dr Darsee told his collaborators that he would do additional experiments at Harvard. The experiments described in the published paper (on 16 dogs) were not done at Emory and according to Dr Braunwald were not done at Harvard. Nevertheless, the additional 16 experiments allegedly performed at Harvard and the revised presentation of results and discussion were adequate for acceptance of the paper for publication. Dr Darsee's colleagues did not know that these experiments were never done.

Many abstracts were submitted for presentation at national meetings without the knowledge of the "collaborators." These abstracts appear to have been submitted after Dr Darsee left Emory University. Much of the work described in the abstracts could not have been done at Emory University, and in fact there is no evidence of it having been done at all. Some of the clinical experiments described in these abstracts would probably have been unacceptable to the Human Investigations Committee of the university. Thus, these abstracts represent fabrication of at least some of the data, and clear dissemblance in that Dr Darsee freely used the names of his former colleagues on the abstracts without notifying them.

There are now allegations that the pattern of fabrication and dissembling reaches back to Dr Darsee's undergraduate days at the University of Notre Dame. Dr Julian R. Pleasants, associate professor of microbiology at Notre Dame, wrote to Dr Braunwald alleging that the two articles Darsee published in the student-run *Notre Dame Science Quarterly* are fabrications. Dr Braunwald sent this committee a copy of Dr Pleasants' letter and copies of Darsee's papers in the *Science Quarterly*. In response to a letter from the committee chairman, Dr Pleasants has sent a detailed analysis of the work of Darsee at Notre Dame, an analysis that makes it unlikely, if not impossible, for Darsee to have done the work described in his papers. . . .

V. Conclusions

The results of the evaluation of Darsee's work by this committee provide overwhelming direct and circumstantial evidence of flagrant and extensive fraud in his research at Emory University and of fabrication of data published in the name of the university even after he was at Harvard. Seven of nine full papers and many of the abstracts must be considered to be invalid, as detailed above.

VI. Recommendations

1. All of the papers and abstracts for which there is no clear evidence that the results are valid should be "retracted." That is, the co-authors should send letters to the editors of the journals, countersigned by Dean Glenn, giving reasons for declaring the data in the papers and abstracts to be invalid. The members of this committee will be available for editorial assistance in composing the letters.
2. The National Institutes of Health should be given a copy of this report.
3. The Georgia affiliate of the American Heart Association should be given a copy of this report.
4. The faculty of the School of Medicine should be informed of the committee's findings. A copy of the report should be made available to members of the faculty who wish to read it.
5. Copies of the report should be sent to Dr Eugene Braunwald and Dean Daniel Tosteson of Harvard Medical School.

6. The scientific and medical community should be informed of the findings of this committee through the news media.

Respectfully submitted,

March 28, 1983

ELMER C. HALL, PH.D.
Professor and chairman, department of biometry

CHARLES M. HUGULEY JR., M.D.
Professor of medicine

PANAGIOTIS N. SYMBAS, M.D.
Professor of surgery

NEIL C. MORAN, M.D.
Charles Howard Candler Professor of pharmacology and chairman of the department;
Chairman of the *ad hoc* committee

ADDENDUM TO REPORT ON RESEARCH OF DR JOHN R. DARSEE AT EMORY UNIVERSITY SCHOOL OF MEDICINE

The report cited two papers that could be considered to be valid, namely paper number 2 ("Neurogenic skeletal myopathy in patients with primary cardiomyopathy." J. R. Darsee, D. O. Nutter, L. C. Hopkins and S. B. Heymsfield. *Circulation*, 59: 492–497, 1979) and paper number 8 ("Hemodynamics of LeVeen shunt pulmonary edema." J. R. Darsee, J. T. Fulenwider, L. F. Rikkers, J. D. Ansley, B. F. Nordlinger, G. Ivey and S. D. Heymsfield. *Annals of Surgery*, 194: 189–192, 1981). Paper number 2 has now been considered to be questionable based on further review by Dr Linton C. Hopkins who has found discrepancies in Table 2 in which data are entered that are not correct, data are entered that were not done and data are entered that Dr Hopkins cannot now confirm. He states also that the labels of Figures 2, 3 and 4 of the paper are incorrect. The patients did exist and the diagnoses are correct. However, because some of the data are incorrect and even possibly fabricated, Dr Hopkins cannot attest to the complete validity of the paper and it is being retracted.

Paper number 8 stands as valid as stated in the report.

Paper number 10 ("Prevalence of mitral valve prolapse in presumably healthy young men." J. R. Darsee, J. R. Mikolich, N.B. Nicoloff and L. E. Lesser, *Circulation*, 59: 619–622, 1979) was considered to be of uncertain validity because of absence of original research data. The co-authors, i.e., Mikolich, Nicoloff and Lesser, however, maintain that the data in the paper are valid although none of them has a copy of the original research results. They state that they did the study and oversaw the writing of the paper. A letter to this point has been sent to the editor of the journal.

Thus, of ten fully published papers, one is valid, one is stated to be valid by the three co-authors and eight are being retracted on the bases described in the report.

Since the report was completed, twelve additional abstracts have been discovered that were not listed on Darsee's bibliography, abstracts that have Emory persons listed as co-authors. Of those additional twelve abstracts, only one can be considered to be valid (number 35 in the list; see below) and that is the one that is related to paper number 10 described above. The other eleven of the recently discovered abstracts were either not known by the co-authors, or were related to some ongoing study. Nevertheless, all eleven should be considered to be invalid or fraudulent. One abstract, for example, ("Improvement in right ventricular function and cardiac output with intrapulmonary artery ballon counterpulsation in patients with acute pulmonary embolism and right ventricular failure." J. R. Darsee, J. R. Mikolich, S. B. Heymsfield. *Chest*, 80: 388, 1981) was published without the co-authors' knowledge and lists the paper as emanating from Harvard Medical School, Boston with no mention of Emory University. The co-authors are not affiliated with Harvard.

Thus of 45 abstracts that are now known only 5 can be considered to be valid. Many of the abstracts appear to be completely fictitious.

NEIL C. MORAN, M.D.
Chairman of committee
May 5, 1983

Medical School Dean Releases Report on Falsification of Research Data

Medical School Dean Daniel C. Tosteson has released a report he received this week from an ad hoc committee he appointed last November to inquire into allegations that a former research fellow, John Darsee, had fabricated research data. The committee also was asked to evaluate the way in which the University handled the matter and to recommend procedures for dealing with future situations involving allegations of fabricated data.

In releasing the report, Tosteson said, "Dishonesty in research strikes at the very heart of true science. It is a contradiction in terms."

Tosteson invited members of the Faculty of Medicine to reflect upon the report and said he would be pleased to receive their comments. He said he intends to use the report "as a basis for developing a process that will guide members of the Faculty should it become necessary to address issues of this kind in the future."

The committee members were: Dr. Richard S. Ross, Dean of the Faculty & Vice Chairman for Medicine at The Johns Hopkins School of Medicine (chairman).

Dr. A. Clifford Barger, Professor of Physiology at Harvard Medical School, Dr. Baruj Benacerraf, Professor of Comparative Pathology and President of Sidney Farber Cancer Institute, Burton S. Dreben, Professor of Philosophy at Harvard, Dr. Saul J. Farber, Professor of Internal Medicine and Chairman of the Department of Medicine at New York University School of Medicine, Gerald Frug, Professor of Law at Harvard; Dr. Robert I. Levy, Dean of the Tufts University School of Medicine, and Dr. Joseph B. Martin, Professor of Neurology at Harvard Medical School.

Here is the full text of the committee's report.

The report is divided into five sections in accordance with the Committee's charge from the Dean:

I. *Chronology.*
The events relating to the charges of dishonesty in scientific research made against Dr. John Darsee are recounted.

II. *Summary of Evidence.*
The evidence for dishonesty in Dr. Darsee's work on three specific research projects is evaluated.

III. *Institutional Response.*
The actions taken by the Harvard Medical School are examined and evaluated.

IV. *Prevention of Dishonesty in Science.*
General observations as to the conditions contributing to dishonesty in scientific research form a basis for suggestions for prevention in the future.

V. *Recommendations to the Harvard Medical School.*

I. Chronology.

Dr. John Darsee came to Dr. Eugene Braunwald's cardiac research laboratory at the Brigham and Women's Hospital (BWH) in July, 1979 with the highest of recommendations from Emory University, where he had received five years of clinical training which included some research experience. During the first 12 to 15 months as a research fellow at Harvard, Dr. Darsee worked closely with the laboratory Director, Dr. Robert Kloner, learning the techniques utilized in the laboratory. He was described as a prodigious worker who learned rapidly, was creative, well read, and excellent with his hands. Because of his extraordinary record and promising future, he was accorded more and more responsibility and given the opportunity to work with less close supervision. Furthermore, on the basis of Dr. Darsee's performance Dr. Braunwald, the Chairman of the Department of Medicine, offered him a faculty appointment as an assistant professor to begin on July 1, 1981.

In the middle of May 1981, some of Dr. Darsee's laboratory colleagues observed gross irregularities in the way he was gathering data and brought this to the attention of Dr. Robert Kloner. Dr. Kloner interviewed Dr. Darsee's colleagues, investigated, and discovered evidence which convinced him that Dr. Darsee had, in fact, fabricated data. On May 22, Dr. Kloner reported his findings to Dr. Eugene Braunwald. Dr. Kloner also confronted Dr. Darsee with these observations on the same day. Dr. Braunwald, conducting an independent investigation, interviewed Dr. Darsee on May 26, the next working day, and told him of the accusations. Dr. Darsee admitted his guilt but was unable to provide a satisfactory explanation for his behavior. The study in which the data fabrication occurred concerned myocardial ischemia in dogs ("Extension vs. Expan-

sion Study"). Dr. Darsee stated that there were no irregularities in any other studies in which he was involved.

Dr. Braunwald considered this to be a problem of the most serious nature, and on May 26, withdrew the offer to Dr. Darsee of a faculty appointment at the Harvard Medical School. On that same date, Dr. Braunwald also informed Dr. Darsee that his National Institute of Health-supported fellowship and his appointment as a research fellow at the Harvard Medical School and at the Brigham and Women's Hospital would expire effective June 30, 1981. Dr. Braunwald informed Dr. Daniel Tosteson, Dean of the Harvard Medical School, and the President of the Brigham and Women's Hospital, that Dr. Darsee had admitted to falsifying research data and that his fellowship and appointments were being terminated. This conversation was followed by a formal letter dated June 6, describing the events that had occurred.

At this point, Drs. Braunwald and Kloner considered it possible that the event in mid-May of 1981 was only a single act of scientific dishonesty, but could not be certain of this without thorough evaluation of all studies in which Dr. Darsee had been involved. Even though Dr. Darsee no longer had BWH and Harvard appointments, his cooperation was necessary in such an investigation, and during the next four months, Drs. Kloner and Braunwald, with Dr. Darsee's assistance, carried out a detailed review and analysis of the work Dr. Darsee had done in their laboratory. This review included consultation with other members of the research team and inspection of primary records. The purposes of the detailed analysis were to determine (1) whether Drs. Kloner and Braunwald could have confidence in other work that Dr. Darsee had done in the laboratory, and whether they could allow him to complete manuscripts covering this work, (2) how to deal with the abstracts submitted to the American Heart Association by Dr. Darsee in May; and (3) whether the fabrication of data in May represented an isolated, impulsive act of an individual working under great pressure. It was felt that careful observation of Dr. Darsee would aid in resolution of these issues.

Close Supervision

One of the studies in which Dr. Darsee had participated during the year prior to May 22, 1981 was an NIH-supported collaborative study of animal models for protection of the ischemic myocardium ("Models Study"). It was decided that Dr. Darsee should complete the experiments in this study under the close supervision of Drs. Braunwald and Kloner. This course of action was selected for two reasons. First, the study was heavily dependent on constant procedures and techniques, and a change of investigator in one of the collaborating laboratories might jeopardize the entire study. Second, the "Models Study" was randomized, blind and carried out simultaneously at four independent laboratories, and the results were analyzed centrally at the NIH. Therefore, any mishandling of data should be readily apparent from a comparison of results from different laboratories.

In October, the work on the collaborative "Models Study" submitted by Dr. Darsee was analyzed along with that from three other laboratories by investigators at all participating institutions including NIH. Inspection of these data by Drs. Kloner and Braunwald revealed a striking lack of variability in the myocardial blood flow measurements for which Dr. Darsee had been responsible and suggested that Dr. Darsee's results should be questioned. Dr. Darsee was asked to produce the primary data upon which the blood flow measurements made prior to May 22 were based, and he was unable to do so. On November 3, the National Heart, Lung and Blood Institute was informed that the laboratory could not take responsibility for any of the work in which Dr. Darsee had participated, and those results should not be included as part of the "Models Study". Also in October, two prospective employers of Dr. Darsee, who made inquiries about him, were told that he had fabricated data in May.

Dr. Darsee had also been collaborating with Dr. Leonard Holman of the Nuclear Medicine Division at the BWH on a study involving the comparison of the measurements of infarct size by different methods. In early November these data collected by Dr. Darsee were critically evaluated and serious discrepancies were discovered. Dr. Darsee was unable to provide an adequate explanation for these discrepancies and because his collaborators no longer had confidence in his results, two manuscripts covering this work were withdrawn.

At this point in November, when it became apparent to Drs. Braunwald and Kloner that it was probable that there was more than one example of irregularity in data collection, all abstracts to the American College of Cardiology meeting from the Harvard laboratories on which Dr. Darsee was the senior author were withdrawn and his presentations to the American Heart Association were cancelled.

Emory Officials Notified

Officials at Emory University where Dr. Darsee had worked previously were notified of the results of Harvard's investigation and advised to conduct a similar audit of all research work with which Dr. Darsee had been associated.

Because of the serious and complex issues raised by this sequence of events, Dean Tosteson decided in mid-November to form an ad hoc committee drawn from senior members of the academic community, both from within and without Harvard University to review the case and develop a full statement of the facts and to recommend procedures for dealing with episodes of this kind in the future.

The first meeting of the committee was held on December 7, 1981, and the second on January 11, 1982. All members were present on January 11 but not on December 7. In addition, on December 24, a subcommittee visited the laboratory where Dr. Darsee worked and reviewed the systems for data collection and preservation used by Dr. Darsee and other members of the laboratory. The entire committee reviewed written reports from all the principals, and Drs. Braunwald and Kloner made oral presentations to the committee at both meetings and were questioned about the events. Dr. Holman of the Division of Nuclear Medicine at BWH appeared before the committee on January 11.

Dr. Darsee met with members of the committee on December 7 and December 24; he appeared before the whole committee (with the exception of Dr. Levy) on January 11. On these several occasions he was questioned extensively by committee members in an attempt to obtain his interpretation of the data brought into question. Members of the subcommittee who visited the laboratory questioned Dr. Darsee and went over his data books with him. The committee is satisfied that Dr. Darsee had adequate opportunity to present his version of the events.

The subcommittee visiting the laboratory and the committee as a whole were convinced that the laboratory directed by Dr. Kloner is well administered with appropriate regard to data collection and review. The visiting subcommittee noted that data books dating back to 1977 revealed an "impressive set of well maintained, annotated and stored raw data that was in good enough condition to allow re-analysis even at this time and seemed complete with protocol, mathematical analysis and finished conclusions". The subcommittee concluded that "the present problem does not appear to be at all referable to the existing Cardiac Research Laboratory standards, policies or procedures nor to overt pressure provided by its Director or Dr. Braunwald. Dr. Robert Kloner and the Cardiac Research Laboratory have maintained an extremely effective system for data collection, analysis and storage".

II. Summary of Evidence.

On the basis of examination of data, and testimony of the senior investigators and Dr. Darsee, the committee concludes that there are serious questions of credibility in three recent studies in which Dr. Darsee was significantly involved.

1. "*Extension vs. Expansion Study*": During May 1981, Dr. Darsee was observed to be fabricating data by producing tracings during a single experiment carried out in a period of a few hours, and labeling them in such a way as to indicate that the observations had been made over a period of two weeks. When confronted with this evidence, Dr. Darsee admitted his wrongdoing, but stated that he was reproducing tracings from previous experiments which had been lost.

2. "*Models Study*": This study involved four laboratories, including that under the direction of Drs. Braunwald and Kloner at the Brigham and Women's Hospital (BWH). Prior to May 22, Dr. Darsee had been responsible for the collection of myocardial blood flow data, which was combined with data collected by others for submission to the NHLBI. Dr. Darsee continued to work on the "Models Study" after May 22, but his work was more closely supervised. There was, therefore, an opportunity to compare Dr. Darsee's pre-May 22 work with his post-May 22 work, and to compare both with the results obtained by others doing the same experiments in three other laboratories.

The myocardial blood flow data submitted by Dr. Darsee prior to May 22 show a variability that all who have examined them judge to be too

small to be real. For example, the standard deviation in Dr. Darsee's pre-May 22 flow measurements was 0.04 to 0.05, as opposed to standard deviations for the comparable observations obtained in other laboratories of 0.28, 0.33 and 0.25. Furthermore, the standard deviation of the flow measurements made by Dr. Darsee after May 22, when he was working under closer scrutiny, was 0.31, an estimate of variability consistent with that obtained in the other three laboratories.

There are also serious discrepancies between the heart weights reported by Dr. Darsee prior to and after May 22. The mean heart weight in studies conducted by Dr. Darsee prior to May 22 was 141.9 gm as opposed to values of 84.8, 87.8, and 71.2 gm from the other three laboratories. The mean heart weight reported by Dr. Darsee for the series of experiments performed after May 22 was 83.1 gm. The average weight of the dogs both before and after May 22 was comparable to those in the other laboratories. Therefore, the unusually high heart weight in Dr. Darsee's pre-May 22 series resulted in a ratio of heart weight (Gm) to dog weight (Kg) which was 7.44 in contrast to values in the other three laboratories of 4.42, 4.36, and 5.07. The ratio in the post-May 22 series from the BWH laboratory was 5.01, a value consistent with that obtained elsewhere.

Dr. Darsee was unable to produce original records to substantiate the pre-May 22 myocardial blood flow measurements. He suggests that the discrepancy between his earlier results and those of others can be explained by his technique. Dr. Darsee further suggests that the discrepancies in the heart weight data are attributable to inaccurate weighing procedures or variation in the treatment of hearts before weighing.

On the basis of careful examination of the evidence, the committee concurred with Drs. Braunwald and Kloner that the data collected prior to May 22, 1981 are clearly beyond the realm of credibility and are suspect as to their origin.

3. "*Creatine Kinase Infarct Size Study*": In this study, the size of a myocardial infarction was determined by different methods. Dr. Leonard Holman determined infarct size by single photon emission computed tomography of the heart in the Nuclear Medicine Laboratory, while Dr. Darsee determined infarct size by means of a method involving analysis of the disappearance of creatine kinase from the blood. The size measurements of these two methods were compared and the correlation obtained was considered to be unbelievably good. Inspection of the original records, where available, indicates that Dr. Darsee's data were incorrectly calculated and reported. The committee concurs in the judgment that the creatine kinase data appear to have been manipulated.

Members of the committee questioned Dr. Darsee at length about all three studies in which his data have been brought into question in hopes that a satisfactory explanation could be found. He admits only to his fabrication of the tracings in the "Extension vs. Expansion Study" in May of 1981, and maintains that all the rest of his work is accurate. He stated to the committee that the data in the "Models Study" obtained before May 22 were accurate and differed from the data from other laboratories because he personally took meticulous care with the methodology. He states that he did not take the same care with the studies after May 22 when he was under great emotional stress and that this lack of care explains the greater variability after May 22 and hence, the discrepancies noted between his pre-May 22 and post May 22 results.

Dr. Darsee ascribes the absence of critical records to the fact that he was burdened with too many records, that storage was a problem, and hence, records were discarded. He states that the records of his calculation in the third study ("Creatine Kinase Infarct Size Study") were not available because the calculations were made on a calculator which gave results only on a visual display and produced no printout.

Examination of all the evidence and testimony forced the committee to conclude that Dr. Darsee did fabricate data on more than one occasion. The committee carefully examined Dr. Darsee's statements and concluded that he did not satisfactorily explain the gross inconsistencies found in his results. He submitted to his superiors and to the scientific community unusual results which are highly suspect and he was negligent in his responsibility to secure and maintain original data to substantiate his alleged results.

III. Institutional Response.

The committee examined the response of the Harvard Medical School to the Darsee incident from the date of discovery on May 22 to the present time. The purpose of this examination was to determine whether any further steps should be taken, whether Dr. Darsee had been treated fairly and what lessons might be learned with regard to the handling of such unfortunate events if they should occur in the future.

The response of the Harvard Medical School should be considered in two phases in accordance with the information available at that time. The first phase response began in late May and early June, following discovery and admission of data falsification by Dr. Darsee. Dr. Darsee's fellowship was terminated on June 30, and his faculty appointment, which had been offered for the following year, was withdrawn. The Dean's Office was notified, and abstracts of work in question were not submitted. Unfortunately, Dr. Holman, Dr. Darsee's collaborator in the Division of Nuclear Medicine did not learn of the fabrication of data in the "Extension vs. Expansion Study" until November and was not informed that Dr. Darsee's appointments had been terminated.

The second phase response occurred in October and November and was initiated by the discovery of the irregularities in the data submitted to the "Models Study" and the discrepancies in the "Creatine Kinase Infarct Size Study". At this point in October and November, when it seemed likely that more than one episode was involved, the granting agency, the NHLBI, was notified and two manuscripts covering the "Creatine Kinase Infarct Size Study" were withdrawn. Dr. Darsee agreed not to present any papers originating from the Harvard laboratories at the annual meeting of the American Heart Association. Officials at Emory University were notified that there were questions about Dr. Darsee's work at Harvard and were advised to examine work performed by him at Emory. In November, the Dean of the Harvard Medical School, Dr. Tosteson, decided that the matter was of such importance that a committee from within and outside Harvard University should be appointed to investigate the situation.

When considered in two phases, the responses of the institution seem reasonable in relation to the information available at the respective times. The first phase response was prompt, but limited, and seemed appropriate for what then appeared to be a single episode of aberrant behavior. The realization that there might be more than one episode of dishonesty appropriately precipitated the larger and more comprehensive second phase response in November and December.

The committee questioned the reason for the delay between the initial event and response in May and June and the fullscale investigation in November. In retrospect, it is clear that Drs. Braunwald and Kloner felt in May and June that they were probably dealing with a single bizarre act by a young man who had performed exceptionally well previously. In the light of this consideration, the plan selected in June seems to be reasonable in that it provided an opportunity to assess the extent of the damage and also to provide for a period of observation under supervision.

As the committee looks back at the events of the summer and fall of 1981, it appears clear that it was helpful to allow Dr. Darsee to continue working in the laboratory because the comparison of the pre-May 22 and post-May 22 data in the "Models Study" proved to be of great value in the final analysis. On the other hand, this course of action was not without cost. Dr. Darsee's presence in the laboratory environment after the events of May proved to be damaging to the morale and productivity of the laboratory.

A systematic review of all work involving Dr. Darsee has been conducted by Drs. Kloner and Braunwald. Primary data have been reviewed and Dr. Darsee's co-workers have been questioned. The committee is convinced that Drs. Braunwald and Kloner have documented the extent of the irregularities and properly concluded that the previously published work from the laboratory in which Dr. Darsee was a member of the group is accurate. None of the work under suspicion has been published. The papers and abstracts containing fabricated data have been withdrawn. This incident has resulted in a strong reaffirmation of the laboratory rules concerning data collection, review and retention.

Darsee Treated Fairly

The committee beheves Dr. Darsee has been treated fairly. For several

months he was dealt with in a manner consistent with his assertion that he was guilty of only one isolated act of data fabrication. It was not until additional information became available in October and November that it became highly probable that this was not the case. The committee approves of the conscientious and responsible attitude of Drs. Braunwald and Kloner in this troublesome affair.

The committee suggests that there are two ways in which the institutional response could have been improved: First, a small committee of senior professors from within the University, but outside the involved department, should have been consulted immediately after the discovery in May. (Perhaps there should be a standing committee to deal with such matters.) In any case, such a committee, be it standing or ad hoc, could have shared the burden with the Dean, the Chairman of the Department, and the Laboratory Director and offered objective advice concerning the management of the problem. The second suggestion has to do with internal communication. In May a systematic search should have been conducted to identify all persons within the institution with whom Dr. Darsee had collaborated, and these persons should have been informed confidentially of the allegations brought against him.

IV. Prevention of Dishonesty in Science.

Dishonesty in scientific research results from interactions between an individual and the environment. Some people are more likely to be dishonest than others, but it is difficult to identify such persons in advance. One factor which may predispose to dishonest behavior would appear to be an intense desire to succeed, a trait generally considered to be laudable. More amenable to modification than personality or self-generated pressures are factors related to the environment. Consequently, the committee believes that the preventive measures most likely to be effective are those directed at the circumstances and practices in the laboratory which stimulate dishonest behavior in some people. In addition, observing the behavior of investigators working in the laboratory can be helpful.

The committee observed that a reported major scientific discovery will be confirmed or refuted rapidly because of the great impact it will have on the direction of science and the strong motivation of other laboratories to verify important data. The probability that spurious or fraudulent results will be discovered by others relates directly to the significance of the report and the number of scientists working in the same field. Therefore, the practice of reporting research results in small segments affords an opportunity for dishonesty because such reports are less likely to be verified by others.

Suggestions Made

The following measures are suggested as helpful in preventing dishonesty in science and its detection if it does occur:

1. Special attention should be given at the time of recruiting to the motivation and integrity of the applicant through careful examination of credentials and claimed accomplishments.

2. Written, detailed, explicit procedures for data gathering, storage and analysis are essential and should be available in all laboratories.

3. Fellows should be supervised by experienced scientists in the laboratory including regular, in-depth scrutiny of the primary data and the calculations leading to the presentation of results. The laboratory director should, by example, develop in the fellows a respect for primary data and its preservation.

4. The conduct of studies which are blind or coded, and the exchange of reagents or methods between laboratories should be encouraged. The repetition of research in the same laboratory and by different research teams is desirable whenever possible.

5. Fellows should be encouraged to work with other colleagues, to share data and to engage in free discussion of results. Secrecy about methods and data should be discouraged.

6. Emphasis should be placed upon the quality and significance of research rather than on quantity and visibility. The laboratory director should accept responsibility for the quality of the work reported from the laboratory and should actively discourage the submission of multiple abstracts to national meetings. Abstracts should be considered as reports of results worthy of publication and as forerunners of complete papers rather than as bids for appearance on a national program.

7. There should be close personal interaction between faculty and fellows. Among the many benefits of

such close interaction could be the early detection of personal problems or unusual personality traits.

Conference Urged

We further suggest that a national conference involving the NIH and universities be convened to consider a number of unanswered questions:

1. Are there steps which could be taken at the national level to encourage the evaluation of young scientists on the basis of quality rather than numbers of publications? These policies must be addressed not only by universities, but also by editorial boards, granting agencies, and study sections. It has been suggested that scientists who fragment their results into many papers should be criticized rather than rewarded.

2. Should national societies examine their policy regarding limitation of multiple abstracts with the same first author?

3. What is the responsibility of the institution discovering dishonesty in scientific research to other institutions, to the scientific and medical community and to the public at large?

V. Recommendations to the Harvard Medical School.

1. There should be a committee, either standing or ad hoc, consisting of senior professors, to which any case of suspected dishonesty in research should be referred *immediately*. The committee would be charged with making a preliminary investigation of any case of alleged dishonesty and recommending an appropriate course of action to the Dean.

2. As soon as an episode of suspected dishonesty is brought to the attention of the supervisor of a laboratory, the episode shall be referred immediately to the department chairman and the Dean. The Dean shall convene and charge the committee. [Naturally, the committee should be sensitive to the effects of their proceedings on the individual under investigation and his/her rights must be scrupulously protected. The sanctions recommended should be appropriate. The committee should also be aware of the profound negative impact a case of dishonesty in scientific research has on the laboratory group in which the event occurred. Every effort should be made to support the innocent collaborators.]

3. As soon as possible the individual under investigation and all persons with whom he/she has collaborated should be informed confidentially as to the nature of the charges and kept informed as the investigation progresses.

4. All work involving that individual should be examined and if the data cannot be substantiated, any publications originating from the work should be withdrawn and the work repeated.

5. A written record of all investigations and deliberations dealing with falsification of results, inadequate documentation, improper manipulation of data, plagiarism and other instances of dishonesty in research should be maintained.

6. Steps 1 through 4 should be completed and a decision reached as swiftly as possible.

Report of the National Heart, Lung, and Blood Institute Special Panel to Review Alleged Misconduct at Brigham and Women's Hospital/Harvard Medical School

Boston, Massachusetts

June 29, 1982

Table of Contents

I. Formation of the Panel - Reasons, Composition 1

II. Charge to the Panel 3

III. Methods and Procedures 4

IV. Purpose and Nature of the Cooperative Study Entitled "Animal Models for Use in the Assessment of Therapy to Protect Ischemic Myocardium" (AMPIM) 7

V. Dr. Darsee's Role in the Cardiac Research Laboratory at the Brigham and Women's Hospital 8

VI. Chronology of Certain Key Events at the Brigham and Women's Hospital 9

A. The AMPIM Study 9

B. Alleged Scientific Misconduct 10

VII. NIH Support Involved 17

VIII. Findings and Conclusions 18

A. Expansion Versus Extension: in vivo Documentation of Infarct Expansion and Shrinkage in the Conscious Dog 18

B. Handling of Data Falsification Incident of May 21, 1981 18

C. Internal Audit Conducted by Dr. Braunwald and Dr. Kloner 19

D. The AMPIM Study at the Brigham and Women's Hospital 20

E. Publications of Dr. John Darsee from Harvard 23

F. Reporting of Regional Myocardial Blood Flow on a Heart That Was Not Analyzed 25

G. Notification of NIH of Non-Renewal of Dr. Darsee's Research Fellowship 26

H. Supervision and Instruction in the Laboratory 27

I. Report of the Ad Hoc Advisory Committee to the Dean of the Harvard Medical School on Dishonesty in Scientific Research 29

J. Current Status of the AMPIM Study 30

K. Summary 31

IX. Recommendations 34

General

Specific

Appendices

1. Summary of Analyses Performed on Data from Brigham and Women's Hospital in the AMPIM Study

2. Summary of Analyses of Data from Selected Publications Coauthored by John R. Darsee, M.D.

3. Report: Examination of Data Books, Tracings, etc. of Dr. John Darsee

4. Report: Dogs Used in AMPIM by the Cardiac Research Laboratory, Brigham and Women's Hospital

5. Report: Use of Animals in Several Published Papers from the Cardiology Research Laboratory, Brigham and Women's Hospital

6. Report: Measurement of Dog Tissue Samples for Labeled Microsphere Activity

7. Report: Measurement of Residual Radioactivity in Additional Samples Received From the Brigham and Women's Hospital (non-AMPIM Cardiac Tissues)

I. Formation of the Panel - Reasons, Composition

In October 1981 the staff of the National Heart, Lung, and Blood Institute (NHLBI) became aware of problems involving the comparability of data reported by investigators in a cooperative study entitled "Animal Models for Use in the Assessment of Therapy to Protect Ischemic Myocardium." This study, referred to by the acronym AMPIM, was launched in fiscal year 1979 with the award of five contracts.

Center	Principal Investigators	Animal Model
Duke University	Keith A. Reimer Robert B. Jennings	Unconscious dog
Brigham and Women's Hospital	Eugene Braunwald Robert A. Kloner	Unconscious dog
Duke University/VA Hospital	Frederick R. Cobb Joseph C. Greenfield	Conscious dog
Johns Hopkins University	Lewis C. Becker William E. Jacobus	Conscious dog
St. Louis University	Jesus E. Val-Mejias Robert Evans	Rat

During the last days of October 1981 it became apparent that some of the data submitted by the investigators of The Brigham and Women's Hospital were different from those submitted by other centers and were not consonant with results obtained in several prior studies. At that time the data in question were regional myocardial blood flows and heart weights. The former showed a remarkable lack of variability and the latter a discordant relationship of left ventricular weight to dog weight. At a meeting of the AMPIM investigators with NHLBI staff on October 30 it was divulged by Dr. Kloner, an investigator at the Brigham laboratory, that a Postdoctoral Research Fellow in this laboratory had been found to have fabricated experimental records and research data in May 1981. In a letter of November 3 Doctors Braunwald and Kloner indicated that they could not vouch or take responsibility for the data provided by this individual.

The episode of observed and admitted data falsification occurred in May 1981 and did not involve the AMPIM study. The individual who committed these acts was John R. Darsee, M.D., who was supported by a fellowship award from the NHLBI (F32-HL5958). The sponsor was Dr. Eugene Braunwald. The award commenced on July 1, 1979 and was to have a three-year duration. However, it was not renewed after June 30, 1981.

Although the senior staff in the Cardiac Research Laboratory at the Brigham and Women's Hospital (Drs. Braunwald and Kloner) had been conducting their own audit of these events, and although the Dean of the Harvard Medical School had on November 25, 1981 established an external review committee, the NHLBI believed that an expert panel commissioned by the granting agency itself was warranted. It seemed possible that several NIH awards for research support might be involved. The integrity and value of the AMPIM Study was in jeopardy. An in-depth review of the circumstances and the corrective actions taken, or yet to be taken, from the standpoints of the NHLBI, the NIH, the public, and the scientific community was judged to be essential. Accordingly, in December 1981, the Institute determined to establish a small panel of senior investigators with extensive experience in cardiovascular research.

The individuals appointed to the Panel were:

Dr. Howard E. Morgan
Evan Pugh Professor of Physiology
Department of Physiology
The Pennsylvania State University
College of Medicine
Hershey, Pennsylvania 17033

Dr. John T. Shepherd
Director for Education
Mayo Foundation
Dean
Mayo Medical School
Rochester, Minnesota 55905

Dr. Francis J. Klocke
Chief
Division of Cardiology
Professor of Medicine and Physiology
State University of New York at Buffalo
Buffalo, New York 14215

Dr. C. Kern Wildenthal
Dean
Southwestern Medical School
The University of Texas Health
Science Center at Dallas
Dallas, Texas 75235

Dr. Morgan served as Chairman of the Panel and Dr. Jerome G. Green, Director of the Division of Extramural Affairs at the NHLBI, served as Executive Secretary for the Panel.

NIH staff members who worked with or observed the Panel were:

Dr. Kent R. Bailey, NHLBI
Dr. David L. DeMets, NHLBI
Mr. Donald C. Gall, Office of the Director, NIH
Mr. Robert B. Lanman, Office of General Counsel
Mr. Joel M. Mangel, Office of General Counsel
Dr. Eugene R. Passamani, NHLBI
Mr. Richard J. Riseberg, Office of General Counsel
Dr. Richard P. Schwarz, NHLBI
Mr. Patrick M. Sullivan, NHLBI

II. Charge to the Panel

In the letter of invitation to the four members of the Panel the Institute's charge to them was outlined. The specific language was:

> The NHLBI Panel is to review circumstances and events in the Cardiac Research Laboratory at the Brigham and Women's Hospital relating to alleged scientific misconduct.
>
> The Panel is asked to:
>
> - describe the nature, extent, and seriousness of the alleged misconduct
> - identify the NIH support instruments and awards that were involved
> - describe and assess the relevant practices, procedures, and supervision that existed, and that now exists, in these laboratories
> - review the reporting and accountability procedures and responsibilities that exist and that were practiced with respect to Brigham and Women's Hospital, Harvard Medical School, and the NHLBI
> - recommend any additional steps and precautions that seem indicated and prudent to deal with the present circumstances or future occurrences and that might serve to prevent such incidents

The members of the Panel were asked to assess overall supervision and instruction of research fellows in the Cardiac Research Laboratory at the Brigham and Women's Hospital during the period 1979 to the present, but not prior to that time. The Panel was not asked to investigate Dr. Darsee's scientific activities prior to his arrival in Boston. Detailed review of published papers from the laboratory was restricted to those which Dr. Darsee had coauthored.

III. Methods and Procedures

The resources and methods utilized by the Panel were diverse; they were designed to be appropriate, equitable, varied, and as rigorous as possible. The methods utilized included the following:

- Letters of inquiry were sent, at various stages, to Dr. Braunwald and his associates at the Brigham and Women's Hospital and the Harvard Medical School. The Panel wishes to note that prompt, detailed, and responsive replies invariably were received.

- The official files of the Institute and the National Institutes of Health were made available including letters, applications, progress reports, and various other types of submitted data.

- Various data tabulations, analyses, and evaluations were made available.

- NHLBI staff accounts, reports, memoranda, and correspondence were made available to the Panel.

- The NHLBI staff prepared, at the Panel's request, several detailed reports, analyses, and evaluations. Important to the Panel's assessments and conclusions were detailed analyses of submitted data from the four AMPIM laboratories using the dog model in terms of measured variables or calculated values plotted sequentially for each animal; this included plots of means and variances over time for each data set and for each center. (Appendix 1).

- Staff prepared analyses of published papers coauthored by Dr. Darsee on work performed at Harvard were also made available to the Panel. (Appendix 2).

- NHLBI expert statistical services and multiple computer-generated analyses were made available to the Panel.

- Books of reports, data tabulations, and correspondence were made available for study in advance and discussed in detail at the Panel meetings.

- Five Panel meetings were held and each provided an opportunity for examination of material, discussion, interviews, and formulation of views, conclusions and recommendations.

- The full report of the Harvard commissioned committee was made available to the Panel by officials at the Harvard Medical School.

- The Panel chairman and executive secretary had an opportunity for direct discussion of these matters with Dr. Richard Ross of Johns Hopkins University, the Chairman of the Harvard Committee. Conversations were held with several other members of this Committee.

- Prior to a planned site visit to Boston, the Panel forwarded specific questions to specified individuals associated with the Brigham and Women's Hospital/Harvard Medical School. Replies were received shortly before the visit along with voluminous and detailed administrative information.

- A two-and-a-half day site visit was conducted at the Brigham and Women's Hospital and Harvard Medical School with an interview schedule arranged in advance. Detailed interviews with Drs. Braunwald, Kloner, and Holman were conducted. (Dr. B. Leonard Holman, Associate Professor of Radiology, had collaborated with Dr. Darsee on two research projects, which were independent of the AMPIM Study; they were conducted in the Division of Nuclear Medicine.) Also interviewed were senior technicians, several current and past research fellows, the Dean of the Harvard Medical School (Dr. Daniel C. Tosteson), the President of Harvard University (Dr. Derek C. Bok), and a senior administrative official at the Brigham and Women's Hospital (Mr. Burton L. Mayblum). These interviews were conducted on an individual and independent basis. The visit also included a tour of the laboratory facilities and an examination of several of Dr. Darsee's laboratory notebooks (for AMPIM and other studies). Additional questions were posed to the senior investigators at the visit.

- Another one-day site visit was conducted by some members of the Panel and NHLBI staff to meet with several key individuals who worked in, or were closely related to, the Cardiac Research Laboratory at the Brigham and Women's Hospital. This visit included interviews with a senior technician, Dr. Braunwald's secretary, Dr. Robert Kloner, and with a former fellow who had been in the laboratory with Dr. Darsee.

- The Panel chairman and two NHLBI staff members spent an additional day examining in detail Dr. Darsee's laboratory notebooks and the laboratory's log book that contained data on use of experimental animals (dogs). (Appendix 3).

- Dr. Darsee agreed to meet with the Panel. He and his attorney met with the Panel and staff at the NIH in Bethesda. A number of questions were forwarded to him in advance of the meeting and additional questions were provided at the meeting. Written responses were received and considered by the Panel. The 3½ hour interview provided Dr. Darsee with a full opportunity to present his own account of events and circumstances.

- At a meeting of the Panel and staff another former fellow in the laboratory, also a contemporary of Dr. Darsee's, was interviewed.

- Several of the individuals interviewed came with carefully prepared, detailed, written opening statements that were read to the Panel and discussed.

- Several telephone interviews were conducted with individuals directly involved with the conduct of the AMPIM Study.

- Drs. Kloner and Braunwald furnished to the Panel tissue blocks from 41 hearts of AMPIM dogs. Radioactivity in these blocks was measured at NIH. (Appendix 6). In addition, cardiac tissue samples from other, non-AMPIM dogs that had been studied in the Cardiac Research Laboratory were forwarded to the NIH for measurement of residual radioactivity. (Appendix 7).

IV. Purpose and Nature of the Cooperative Study Entitled "Animal Models for Use in the Assessment of Therapy to Protect Ischemic Myocardium" (AMPIM)

In January 1978, a workshop was conducted at the National Heart, Lung, and Blood Institute to examine and discuss scientific advances in the field of experimental myocardial infarction. The focus of this workshop was animal models which might be useful for the assessment of the efficacy of therapeutic agents aimed at limiting infarct size. The workshop concluded that there were many useful models and numerous interventions with potential efficacy to limit infarct size; however, the wide variation in methods and procedures among laboratories made the comparison of experimental results difficult. Accordingly, the workshop recommended that several animal models be characterized and developed so that there would be a more effective means for comparison of experimental results among laboratories.

Upon the recommendation of the Cardiology Advisory Committee in April 1978 and the National Heart, Lung, and Blood Advisory Council in November 1978, a request was issued on February 15, 1979 for proposals (RFP-NHLBI-79-20) which would characterize and develop the unconscious dog model, the conscious dog model, and the rat model of experimental myocardial infarction. Fourteen proposals were received in response to this solicitation, and after external peer review, staff review and site visits, five contracts were awarded in September 1979.

Dr. Robert Jennings of Duke University was selected by the NHLBI Program Office to be the chairman for the study. Richard P. Schwarz, Jr., Ph.D., was project officer for the program and Kent R. Bailey, Ph.D., was the NHLBI statistician. The NHLBI served as the data coordinating center for the study.

V. Dr. Darsee's Role in the Cardiac Research Laboratory at the Brigham and Women's Hospital

John R. Darsee, M.D., joined the Cardiac Research Laboratory at the Brigham as a postdoctoral fellow in July 1979, having received superlative recommendations from his preceptors at Emory University. His research productivity at the Brigham was extraordinary. During the first 15 months he was first author on five papers submitted to quality scientific journals and subsequently published; these described a variety of dog experiments dealing with the limitation of infarct size with therapeutic agents. Appendix 2 lists these papers.

During the tenure of his fellowship Dr. Darsee collaborated with Dr. B. Leonard Holman, Associate Professor of Radiology in the Division of Nuclear Medicine. Their studies dealt with the quantification of myocardial infarcts in animals and humans measured by emission computed tomographic techniques and by standard laboratory methods (serum creatine kinase enzyme activity).

Dr. Darsee was also involved in several other research projects.

On September 9, 1980, Dr. Darsee began to perform the major responsibilities he had been assigned for the AMPIM Study. These were the performance of the surgery and occlusion of the coronary artery; acquisition of the hemodynamic data, including regional myocardial blood flow; excision, sectioning and weighing of the hearts; determination of the size of the ischemic bed ("risk region"); and filling out the data forms for all 50 dogs randomized between September 9, 1980 and July 25, 1981. At the request of Dr. Kloner, Dr. Darsee was added to the AMPIM Coordinating Committee in December 1980.

VI. Chronology of Certain Key Events at the Brigham and Women's Hospital

A. The AMPIM Study

The AMPIM Study was divided chronologically into three phases: Protocol Development (Phase I), Performance of Randomized, Blinded Studies (Phase II), and Analysis of Data (Phase III).

Phase I (October 1979 - June 1980) - Protocol Development

In this phase, the investigators met on five occasions during which they reported data and held discussions concerned with the development of a uniform protocol. Since the rat model at St. Louis University was so different from the conscious and unconscious dog models, the investigators at this institution elected to pursue their studies independently, but concurrent with those in dogs in the four other laboratories.

The protocol developed for the unconscious and conscious dogs included uniform procedures among the four laboratories for animal selection, anesthesia, surgery, coronary occlusion, sacrifice, sectioning of the ventricle and quantification of infarct size. The methods for recording hemodynamic data and regional myocardial blood flow were standardized, as was the method for assessment of the size of the ischemic bed ("risk region"). Two drugs were selected for investigation: Verapamil (a calcium entry blocker) and Ibuprofen (a non-steroidal anti-inflammatory agent). Dosage and route of administration of the drugs were standardized, and the dogs were to be randomized to one of three treatment groups:

Group I	Verapamil
Group II	Ibuprofen
Group III	Ibuprofen Vehicle (control group)

Group I was to be performed single-blind, with the person performing the surgery being aware that the animal was assigned to the verapamil group; however, this was not known until 5 minutes prior to the commencement of drug infusion. Groups II and III were to be performed double-blind, with the laboratory personnel having no knowledge of which treatment the animal was receiving. It was agreed that all post-mortem analyses were to be performed by individuals who had no knowledge of the experimental group to which the dog was assigned.

The randomization scheme and the coding and labeling system for the drugs were developed and coordinated by the NHLBI through the US Public Health Service Dispensary at Perry Point, Maryland.

Phase II (July 1980 - July 1981) - Performance of Randomized, Blinded Studies

Commencement of Phase II began in July 1980, with the four laboratories performing the experiments according to the standardized protocol. In the twelve months that followed, the four laboratories randomized a total of 180 dogs (51 at Duke, 50

at Brigham and Women's, 38 at Johns Hopkins and 41 at Duke/VA Medical Center). Approximately 180 data points were reported for each dog.

The blinded data were accumulated at each laboratory and reported to the NHLBI on a uniform data reporting form. When received at the NHLBI, the data were checked for appropriate format and entered into an SAS (Statistical Analysis System) computer data base for analysis.

To ensure that the laboratories were adhering to the protocol, it had been decided to perform an interim analysis of the data, at approximately the mid-point of the study, without revealing to the investigators the likelihood that either drug was effective in limiting infarct size. This meeting was held on April 9, 1981, in Bethesda, Maryland. At the meeting, data analyses for the unconscious and conscious dog studies were presented.

Data for 81 randomized dogs were available and revealed no definitive evidence of efficacy with either drug in either model. There was a slight lack of comparability between the Duke and Brigham and Women's control groups in the unconscious model preparation. It was decided to complete the studies as originally planned, and to meet in early fall 1981 for discussion of the final data.

Phase III (July 1981 -) - Analysis of Data

Data for an additional 99 randomized dogs from the four laboratories were furnished to the NHLBI during July - September, 1981. The data were incorporated into the computer data base and tabular data sets were generated, giving the complete data set for each of the four laboratories.

B. Alleged Scientific Misconduct

The following is a summary of certain key events for the period May 15, 1981, when certain irregularities concerning Dr. Darsee were detected and reported at the Brigham and Women's Hospital, through June 29, 1982, when the report of this Panel was completed.

Date	Summary of Events
May 15, 1981	A Research Fellow and a technician in the laboratory expressed suspicions concerning data irregularities in some experiments included in Dr. Darsee's unpublished abstract on infarct extension/expansion. Dr. Robert Kloner, the laboratory chief to whom these concerns were expressed, investigated these allegations. He found that a photocopy of tracings on continuous undated strip paper had been labeled 24, 48, and 72 hours and 7

days. Dr. Kloner concluded that it was "highly unlikely" that the recording apparatus was not used for other studies during such a long period. Dr. Kloner informed Dr. Darsee that he wished to examine all primary tracings for this study.

May 21, 1981

Dr. Darsee was observed performing a dog experiment for his infarct extension/expansion study. He was observed by three individuals to be labeling the tracings he obtained that day, from the same dog, "24 hours", "72 hours", "one week" and "two weeks." Dr. Kloner was informed of this. On the next day Dr. Darsee presented these tracings to Dr. Kloner as valid data to be included in his abstract. When confronted by Dr. Kloner with the allegations of data fabrication and the mislabeled borders of the strip charts which Dr. Kloner had recovered from the trash, Dr. Darsee admitted his guilt and apologized. Dr. Kloner kept Dr. Braunwald informed of these circumstances.

May 26-29, 1981

Dr. Braunwald conducted his own investigation of the matter. After obtaining an admission of guilt from Dr. Darsee, he withdrew a prior offer of a Harvard faculty appointment which had been made to Dr. Darsee. Dr. Braunwald and Dr. Kloner began an audit of all of Dr. Darsee's data and publications from the Harvard Medical School/Brigham and Women's Hospital Cardiac Research Laboratory. Dr. Braunwald informed Dr. Daniel Tosteson (Dean, Harvard Medical School) orally of the events.

June 6, 1981

Dr. Braunwald informed Dr. Tosteson, Dr. Adelstein (Dean for Academic Affairs) and Dr. Petersdorf (then President of the Brigham and Women's Hospital) in writing of the events.

May 26-July 25, 1981

Dr. Darsee was allowed to continue to perform work on the AMPIM Study, but under somewhat closer supervision than previously.

June 27, 1981

Dr. Darsee wrote to NIH and indicated he would not be requesting the third year of his fellowship award.

June 30, 1981

In accordance with a June 25 request from Dr. Braunwald, an institutional official at Brigham and Women's Hospital (Mr. B. Mayblum) formally notified NIH by letter that Dr. Darsee's fellowship was to terminate on that day.

October 5, 1981

NHLBI staff completed the preparation of the AMPIM Study data sets. Discussions revealed that there was a comparability problem with respect to the endpoint

data (ratio of infarct size to the size of the left ventricle) in both the conscious and unconscious dog models.

October 6, 1981 — The NHLBI staff (Dr. Richard Schwarz) called the Chairman of the study (Dr. Robert Jennings) and the key investigators in all four laboratories to inform them that there was a comparability problem with regard to the endpoint data. It was explicitly stated that there was no explanation for the differences among the endpoint data, and that the Coordinating Committee would have to attempt to identify the reasons for the differences. Plans were made to meet, as previously arranged, in Bethesda on October 30 to discuss the data. The investigators were informed that they should expect to receive all the data within a week.

October 9, 1981 — The complete set of unblinded data was mailed to the four laboratories.

October 20, 1981 — A second mailing was sent to the four laboratories, providing additional data which would assist the investigators in the attempt to determine the cause(s) of the endpoint data discrepancies.

October 21, 1981 — As the result of discussions with Dr. Robert Jennings and other AMPIM investigators, concerns were raised to NHLBI staff that the problem of data comparability might not be restricted to the endpoint data alone. Specifically, some of the investigators expressed concern that the heart weight and myocardial blood flow data from the Brigham and Women's laboratory were inconsistent with the data from the other laboratories. The reasons for the inconsistency were unclear.

October 23-29, 1981 — On October 23 the NHLBI staff at Dr. Keith Reimer's and Dr. Robert Jennings' suggestion, generated graphs of the ratio of infarct size to the size of the left ventricle vs subepicardial blood flow and a graph of left ventricle weight vs animal weight for each center; this was done with the intent of distributing them for discussion at a meeting of AMPIM co-investigators on October 30. Two principal observations were noted. First, the flow data for the Duke, Johns Hopkins, and Duke/VA laboratories showed an approximate inverse relationship between (infarct size/left ventricle) and subepicardial flow for all three treatment groups; however, this relationship was lacking in the Brigham data. Second, the ratio of left ventricle weight to dog weight was much higher for the

	Brigham data than for any other laboratory. NHLBI staff were very concerned about these problems of data comparability since there was evidence that the problem was not restricted to the endpoint data alone. Dr. Kloner was notified of the new areas of concerns by October 29.
October 29, 1981	Dr. Kloner called the NHLBI early in the day to inform staff of a serious matter concerning the integrity of the Brigham and Women's data in the AMPIM Study: Dr. John Darsee had admitted to a single instance of data fabrication in another study in May 1981.
October 30, 1981	The AMPIM Coordinating Committee met in Bethesda. After review and discussion of the data, Dr. Kloner told the group about the incident involving Dr. Darsee and that the Brigham laboratory could not vouch for all of the data they had submitted.
November 3, 1981	Dr. Braunwald and Dr. Kloner submitted a letter to NHLBI concerning Dr. Darsee's prior admitted misconduct and the AMPIM data irregularities. They indicated that the Brigham laboratory could not vouch for some of the data collected by Dr. Darsee that they had submitted. They offered to repeat the AMPIM Study. Discrepancies between pre-May 22 and post-May 22 results in Dr. Darsee's blood flow and heart weight data, which had been identified by Dr. Kloner late on October 29, were under continuing analysis in Boston and were called to NHLBI's attention in Dr. Braunwald's letter to Dr. Green of January 12, 1982.
November 1981	Dr. Braunwald was informed of Dr. Darsee's continuing collaboration with Dr. Leonard Holman in the Department of Nuclear Medicine; he reviewed the manuscripts that they had submitted for publication and discovered data irregularities. These irregularities were discussed with Dr. Holman who withdrew these papers.
November 25, 1981	Dean Tosteson established an Ad Hoc Advisory Committee to the Dean of the Harvard Medical School on Dishonesty in Scientific Research.
December 1981	The following incident was reported to Dr. Kloner who, in turn, reported it to Dr. Braunwald: Dr. Darsee was observed by a Research Fellow and a technician on May 14, 1981 as he performed an experiment on a dog in his "Pacing Study." The protocol was not completed and Dr. Darsee placed the carcass, with heart <u>in situ</u>, into a disposal bag. This

seemed unusual because, in the Cardiac Research Laboratory, standard procedures for determination of regional myocardial blood flow require the excision of the animal's heart. The technician requested a Research Fellow to remove the heart and the technician retained it. Several months later (in December), staff members in the laboratory had an opportunity to review Dr. Darsee's data notebooks and noted that blood flow values were recorded for that particular experiment.

January 25, 1982 — The report of the Ad Hoc Advisory Committee to the Dean of the Harvard Medical School on Dishonesty in Scientific Research was released. No additional evidence of scientific misconduct was detected.

February 17, 1982 — The first meeting of the National Heart, Lung, and Blood Institute Special Panel to Review Alleged Misconduct at the Brigham and Women's Hospital/Harvard Medical School took place. Data irregularities in the AMPIM Study involving parameters other than blood flow and heart weight were reported to the Panel. An analysis of all AMPIM data was begun. These analyses revealed irregularities in all but one of the parameters that were measured at the Brigham and Women's Hospital.

Irregularities in data reported in papers in which Dr. Darsee was a coauthor originating from the Brigham and Women's Hospital were detected and an analysis of these papers begun.

February 28, 1982 — In written responses to questions posed by the NHLBI Panel, Drs. Braunwald and Kloner stated that they were suspicious of the validity of the massive reduction in infarct size from ibuprofen and verapamil in the AMPIM results provided by Dr. Darsee. These reductions were much greater than those found in earlier studies in the Cardiac Research Laboratory.

March 2-4, 1982 — The Panel and NHLBI staff conducted a site visit to the Brigham and Women's Hospital and the Harvard Medical School. The Panel interviewed current and former staff members and visited the Cardiac Research Laboratory. During this site visit there was discussion of data irregularities in the AMPIM Study and in Dr. Darsee's published papers; there was also a review of data books in the laboratory and a study of Dr. Darsee's use of dogs in multiple publications

was undertaken. The Panel was informed that Dean Tosteson, with the concurrence of Dr. Braunwald, had appointed senior faculty members to reopen Harvard's investigation of Dr. Darsee's published work from the laboratory and to review his other previous experiments.

March 13, 1982 — Dr. Darsee submitted to the Panel a proposal that would allow him to repeat the Brigham and Women's Hospital portion of the AMPIM Study.

March 18, 1982 — Because of the widespread data irregularities in the AMPIM Study, a careful review of the dog log book was undertaken. When the dates of experiments on AMPIM dogs recorded in the log were compared with those on the randomization cards, many discrepancies or omissions were noted. Also reviewed at this time were the data books for published papers emanating from the Brigham laboratory in which Dr. Darsee was a coauthor. Subsequently, NHLBI staff prepared, with Dr. Kloner's assistance, an analysis of the use of dogs in these papers.

March 30-31, 1982 — The Panel was informed that Dr. Darsee had reported regional myocardial blood flow on a heart that he had not excised. Analysis of the use of dogs in published papers coauthored by Dr. Darsee revealed misrepresentation of the randomization methods as stated in the papers.

April 12, 1982 — Two members of the Panel and the Executive Secretary visited the Brigham and Women's Hospital to gather information concerning randomization procedures in the laboratory and the reported measurement of regional myocardial blood flow in the dog with heart left in situ.

April 27, 1982 — The Panel was informed by Dr. Braunwald and Dr. Kloner that many of the dogs on which Dr. Darsee had reported regional myocardial blood flow were not included in the radioactive disposal log, including the first 35 dogs used in the AMPIM Study. From dog 36 onward in the AMPIM Study, the dogs were entered in the log.

May 31, 1982 — Dr. Braunwald reported that Dr. Kloner had analyzed residual radioactivity in histologic tissue blocks from the AMPIM Study and found that blocks from dogs prior to May 22 contained little if any radioactivity while blocks from dogs after May 22 still exhibited substantial radioactivity.

June 2-3, 1982 — Tissue blocks from hearts of dogs used in the AMPIM Study were forwarded to NIH for analysis of residual radioactivity. The NIH analysis confirmed that of Dr. Kloner.

June 6-7, 1982	The last meeting of the NHLBI Special Panel took place and a final report was prepared. Prior to the meeting, copies of the Panel's draft report were sent for comment to Drs. Braunwald, Darsee, Kloner, and Tosteson. Their written comments were reviewed in detail and taken into account in the preparation of the final report.
June 17, 1982	Drs. Kloner and Braunwald forwarded to the NIH additional cardiac tissue samples from ten non-AMPIM dogs studied by Dr. Darsee. These tissues were measured for residual radioactivity. Although isotope-based blood flow data had been reported (and, in some instances, published) for these animals, no residual radioactivity was detected.

VII. NIH Support Involved

The NIH grants, contracts, and awards that Dr. Darsee was involved with or to which he related directly are listed below. All were awarded by the National Heart, Lung, and Blood Institute.

Postdoctoral Research Fellowship

F32-HL05958 — NRSA individual fellowship
Three years committed support
Year-01: 7/1/79-6/30/80
Year-02: 7/1/80-6/30/81
Year-03: NOT ACTIVATED

Research Project Grant

R01-HL23140 — Determinants of Irreversible Ischemic Injury in Heart
Project Period: 6/30/78-6/30/85
Dr. Darsee's involvement: 7/1/79-6/30/80

Specialized Center of Research (SCOR) Grant

P50-HL26215 — Ischemic Heart Disease
Project Period: 1/1/80-12/31/84
Dr. Darsee's involvement: 7/1/80-6/30/81

Research Contract

N01-HV-92921 — AMPIM Contract
Contract Period: 9/30/79-6/30/82
Dr. Darsee's involvement: 7/1/80-7/25/81

VIII. Findings and Conclusions

A. Expansion Versus Extension: In vivo Documentation of Infarct Expansion and Shrinkage in the Conscious Dog

The initial detection of data falsification by Dr. John Darsee involved this study.

1. Findings: Suspicions that data in this study were fabricated were raised by a laboratory technician, and two Research Fellows. Dr. Kloner was informed of these suspicions. He uncovered apparently mislabeled strip chart recordings and requested Dr. Darsee to produce the raw data and tracings from the pertinent chronic dog experiments. Soon after this request, on May 21, Dr. Darsee was observed to label tracings obtained within a few hours as having been recorded over two weeks. Dr. Darsee admitted falsification of data when confronted with the evidence by Dr. Kloner and Dr. Braunwald.

2. Conclusions: Responsibility for data falsification in the expansion versus extension study rests with Dr. Darsee and is acknowledged.

B. Handling of Data Falsification Incident of May 21, 1981

1. Findings: As soon as the incident was verified, Dr. Kloner notified Dr. Braunwald who acted promptly to notify Dean Tosteson and officials of the Brigham and Women's Hospital. Dr. Braunwald believed that this was an isolated event. He did not notify the NIH and felt that there was no legal responsibility to do so. In his statement to the Panel in early March, he also expressed concern that a confidential channel for notifying NIH was not available, and noted that there were no NIH guidelines or precedents implying that he should do so under the circumstances then believed to be present (i.e., that Dr. Darsee had engaged in only one isolated incident of misconduct). Dr. Braunwald withdrew his offer to Dr. Darsee of an Assistant Professor's appointment at the Harvard Medical School (Beth Israel Hospital) and Dr. Darsee's NIH fellowship was terminated as of June 30. Dr. Braunwald retained Dr. Darsee in his laboratory to complete the AMPIM Study and two other projects in which he had been involved. Neither the Chairman of the AMPIM Study, the collaborating AMPIM investigators, nor the NHLBI were consulted in the decision to retain Dr. Darsee in the AMPIM Study. With Dean Tosteson's approval, Drs. Braunwald and Kloner took responsibility for scrutinizing all of the data Dr. Darsee had collected during his tenure in the Harvard laboratory. Priority was given to published manuscripts, followed by papers in press, abstracts to be published, and finally ongoing work.

2. Conclusions: Drs. Braunwald and Kloner acted in a prompt and responsible manner in notifying Dean Tosteson of Dr. Darsee's data falsification. Dr. Braunwald has stated that his concerns at that time were, "first, and most important, a deep obligation to the scientific community to ensure the integrity of all research and data emanating from his research laboratory; second, a concern that Dr. Darsee be treated fairly as both a professional and as a person; third, a duty to meet fully his legal obligations arising

from the situation, including his legal obligation to the NIH, Harvard University and the Brigham and Women's Hospital." While sensitive to these concerns, as well as the altered perspective afforded by retrospective view, the Panel believes that there were also responsibilities to co-investigators in the AMPIM Study and the National Heart, Lung, and Blood Institute that were not adequately discharged. Although the initially discovered instance of data falsification has been said not to involve a federally supported study, the Panel regards all of Dr. Darsee's scientific effort between July 1979 and June 30, 1981 as involving public funds since he was a full-time NHLBI-supported Research Fellow. Although NIH has not issued guidelines concerning the responsibility of grantees, sponsors, or contractors to report actual or alleged misconduct, the Panel believes that NHLBI should have been informed of the original incident promptly. In addition, collaborating AMPIM investigators - or at least the Chairman of the AMPIM Study - and NHLBI should have been offered an opportunity to comment on the proposed continuation of Dr. Darsee in a study in which they as well as the Boston group were importantly involved. Limiting the investigation of Dr. Darsee's work to an internal audit conducted by Drs. Braunwald and Kloner may have contributed to the delay in discovering the extent of the data falsification. Dean Tosteson should have implemented an institutional audit conducted by faculty members from outside the involved Department and laboratory, and ensured that the auditing group was responsible for a complete and timely report. Such a course could have resolved the situation at an appreciably earlier time.

C. Internal Audit Conducted by Dr. Braunwald and Dr. Kloner

1. Findings: The initial audit (May - October 1981) detected no further instances of wrongdoing. Following the NHLBI staff's October disclosure of noncomparable results in the AMPIM Study, additional evidence of wrongdoing was uncovered. Results of the internal audit were summarized by Drs. Braunwald and Kloner in a letter to Dean Tosteson on December 29, 1981. "We have re-reviewed the results of these experiments and other than his participation in the 'Models'* collaborative study (as well as the original experiment which raised this issue in the first place) have confidence in the results of these other studies. One of us (EB) has expressed serious concern about the data in one of Dr. Darsee's experiments carried out in Dr. Holman's laboratory. We then stepped back and have attempted to analyze whether the principal conclusions from his work in our laboratory had been verified by other investigators in our laboratory working totally independently of Dr. Darsee...Thus, work carried out independent of Dr. Darsee by other investigators in our laboratory, before, during, and after Dr. Darsee's experiments, corroborate the major conclusions that were reached in studies in which Dr. Darsee has participated. We believe that these considerations strengthen our conclusion (based on a detailed analysis of his role in these studies and of the results of individual experiments) that no misleading results have been released from our laboratory."

*'Models' Study refers to the AMPIM Study.

63-471 - 87 - 10

As will be described subsequently, the Panel's findings extend these conclusions in several respects and differ from them in others: (1) All measurements performed by Dr. Darsee in the AMPIM Study prior to May 22 show disturbingly little variability; (2) During the conduct of the AMPIM Study, there was an important deviation from the agreed-upon protocol in that Dr. Darsee was allowed to perform all but one of the post-mortem analyses in an unblinded manner; (3) The published papers from the Cardiac Research Laboratory in which Dr. Darsee was first author contain statistical aberrations which cast doubts upon the primary data; (4) The raw data needed to verify the reported values of blood flow and hemodynamics in these papers rarely are available (Appendix 3); (5) These papers include descriptions of "randomization" of animals which are misleading as to the actual process by which animals were selected for control and experimental protocols; (6) Not all relevant allegations by co-workers in the laboratory were adequately investigated in 1981.

The Panel's findings in regard to the experiment carried out in Dr. Holman's laboratory are similar to those expressed by Dr. Braunwald. The study in question involved a comparison of infarct size measured by creatine kinase techniques with infarct size measured by single photon emission tomography. Creatine kinase activity-time curves were constructed retrospectively from routine clinical measurements; infarct size calculations employed variable values for the rate constant of disappearance of creatine kinase. The determinations of infarct size using creatine kinase were the responsibility of Dr. Darsee; his calculations are not available for review. The reported agreement of the creatine kinase and tomographic infarct size measurements was considered unrealistically good by Dr. Braunwald, an expert he consulted, and Dr. Holman. The paper reporting these results was withdrawn prior to publication.

2. Conclusions: The Panel believes that the internal audit conducted in 1981 by Drs. Braunwald and Kloner was insufficiently rigorous and not definitive. It should have concentrated on whether data reported were internally consistent and verifiable in retrospect, and on whether experiments had been performed as described. As mentioned above, an audit should have been conducted by faculty members from outside the involved laboratory and Department. A request to NHLBI to assist in an examination of AMPIM data would also have been prudent. The Panel agrees that the creatin kinase data generated by Dr. Darsee in the study with Dr. Holman cannot be believed.

D. The AMPIM Study at Brigham and Women's Hospital

Six areas of concern were identified by the Panel:

1. Data irregularities prior to May 22 (dogs 1-35)

a. Findings: In addition to the irregularities in blood flow and heart weight data noted in late October, minimal variability is present in many other measurements carried out by Dr. Darsee (heart rate, blood pressure, the amount of additional anesthesia required during a study, and post-mortem measurements

of risk region based on intracoronary dye injections). (Appendix 1). Raw data are not available for verification. Dr. Darsee carried out these experiments on his own. Dr. Kloner occasionally visited the room in which experiments were performed, but took an active role in the conduct of only a few experiments. Dr. Braunwald was seldom in the laboratory and had no direct supervisory role.

b. Conclusions: Dr. Darsee was responsible for generation of data with unbelievably low variability. His claim that the low variability in the data prior to May 22 was due to his extreme care in the conduct of these experiments is unconvincing.

2. Adherence to the protocol agreed to by all investigators in the study

a. Findings: A departure from the written protocol occurred in all dog experiments (1 through 50). The same person (Dr. Darsee) carried out the surgery, randomized the dogs, and did all but one of the post-mortem analyses in an unblinded fashion. The protocol is as follows:

> "Blinding of Post-Mortem Analyses - All tissue samples, blood flow data and other data associated with the post-mortem analyses will be labeled with the animal ID number only. No indication of the experimental group to which the animal has been assigned will be given to the person who performs the post-mortem analyses."

b. Conclusions: The ultimate responsibility for this departure from the protocol must rest with Dr. Braunwald, the principal investigator. Drs. Kloner and Darsee must also be held responsible, because as co-investigators both attended meetings of the AMPIM investigators and should have been thoroughly familiar with the provisions of the protocol. (See Section J, page 30).

3. Supervision of Dr. Darsee after May 22 (dogs 36-50)

a. Findings: After Dr. Darsee's initial falsification of data had been detected and admitted, Dr. Kloner was more frequently in the room in which AMPIM studies were performed, but Dr. Darsee's specific assignments were not changed. Thus, Dr. Darsee was left to carry out the remaining 15 dog experiments under conditions that continued to deviate from the AMPIM protocol. Appropriate records were retained on these dogs (Appendix 3) and the pattern of variability of most measurements became similar to that reported in other AMPIM laboratories.

b. Conclusions: The advisability of allowing Dr. Darsee to continue to do AMPIM experiments after May 22 must be questioned. Allowing him to do these later experiments with limited supervision, in continued deviation from the AMPIM protocol, and without consulting AMPIM co-investigators, was a mistake.

4. Other measurements in the AMPIM Study

a. Findings: The only measurement in the Brigham and Women's AMPIM data that showed a consistent degree of variability throughout the study, similar to that

in other centers, was the histological measurement of infarct size expressed as a percentage of histologically determined risk region. This measurement was the responsibility of Dr. Kloner and two histology technicians. The technicians were instructed in the conduct of the measurements, but carried out most of them on their own. When they felt they were having difficulty with specific specimens, they consulted Dr. Kloner. Some gross examinations of the areas planimetered were performed by Dr. Kloner.

Large favorable effects of ibuprofen and verapamil on infarct size were observed in the Brigham and Women's AMPIM data. These data were dependent on a combination of the measurements made by Dr. Kloner and his technicians and of separate measurements made by Dr. Darsee. These observed favorable effects are now felt by Drs. Braunwald and Kloner to be larger than those previously noted in their own or other laboratories.

b. Conclusions: There is no evidence that any irregularities occurred in measurements made by anyone other than Dr. Darsee. The Panel believes that Dr. Kloner should have been regularly and more directly involved in the histological measurements which represented the primary endpoint of the study, but there is no evidence that the performance of the technicians was not satisfactory.

When considered in the context of the extensive data irregularities, the remarkably large favorable ibuprofen effects obtained at Brigham and Women's Hospital raise the possibility that the "blinding" of ibuprofen and ibuprofen-vehicle was broken. If so, in view of Dr. Darsee's duties in the study, it is not apparent how anyone other than he could have been in a position to accomplish this. The Panel concludes that the large favorable ibuprofen and verapamil effects on infarct size cannot be relied upon.

5. Accounting of dogs used in the AMPIM Study

a. Findings: Of the first 35 dogs used in the AMPIM Study, only 15 were entered in the laboratory log (Appendix 4). Of these 15 dogs, the randomization date often did not match the date of entry into the laboratory log book. The laboratory also maintained a separate log to record the disposal of dogs that had received injections of radioactive materials. With respect to this latter book, Dr. Braunwald indicated to the Panel that "the check revealed that all Models dogs from #36 on (the first Model's dog after May 26), were recorded on the list, but none from Nos. 1-35 were so recorded."

b. Conclusions: Dr. Darsee was responsible for entering dogs in the laboratory log and for randomization and must be held primarily accountable for these instances of inadequate record keeping. Poor record keeping raises the possibility of assignment of animals to control or treatment groups after experiments had been completed and some of the results were known. The absence of dogs #1-35 in the radioactive log raises doubts as to whether these dogs received injections of radioactive microspheres for the measurement of myocardial blood flow.

6. Residual radioactivity in AMPIM cardiac tissue

a. Findings: As noted above, Drs. Braunwald and Kloner detected, and called to the Panel's attention, the fact that AMPIM dogs #1 through 35 are not listed in the radioactive dog disposal log; dogs studied after May 22 (#36-50) were listed in the log. Dr. Kloner retrieved histologic tissue blocks on a number of AMPIM dogs from the Cardiac Pathology Laboratory and residual radioactivity was determined. He reported to the Panel that in seven dogs studied just prior to May 22 the number of counts in the tissue blocks was indistinguishable from the background level; in seven tissue blocks from dogs studied by Dr. Darsee immediately after May 22, the radioactivity level was much higher than background.

Per agreement with the NHLBI staff, Dr. Kloner forwarded to NIH two histology blocks for each of the 41 AMPIM dogs that survived the complete protocol. These blocks were counted in an especially sensitive radioactive detector. Tissue from dogs studied prior to May 22, 1981 had radioactivity levels similar to background levels; tissue from dogs #36 through 50 had significantly greater radioactivity. The report from the NIH Nuclear Medicine Department (Appendix 6) reveals a clear and marked difference between these two periods.

b. Conclusions: The radioactive dog disposal log was the responsibility of two technicians in the Cardiac Research Laboratory. No other staff members made entries into this log. Dr. Darsee contends that the log book is grossly inaccurate and that, as required by the protocol, he injected radioactive microspheres into each of the AMPIM dogs. The residual radioactivity determinations performed at two separate facilities refute this contention.

Although Dr. Darsee submitted regional myocardial blood flow data derived from techniques utilizing radioactive materials, the Panel believes that the evidence indicates that Dr. Darsee did not inject radioactive microspheres into AMPIM dogs before May 22, 1981. The Panel notes also that the residual radioactivity counts are entirely consistent with the entries in the radioactive dog disposal log and with the statistical analyses noted above.

Drs. Braunwald and Kloner deserve full credit for initiating the measurement of residual radioactivity in these cardiac tissues. This technique led to the important discovery that there was an absence of residual radioactivity in many of the hearts in which Dr. Darsee purported to have measured myocardial blood flow by injecting radioactive microspheres.

E. Publications of Dr. John Darsee from Harvard

1. Analyses of tabular data

a. Findings: Five publications from the Harvard laboratory in which Dr. Darsee was the first author were evaluated in view of the evolving findings indicating more widespread data irregularities than originally appreciated. (Appendix 2). These studies were supported by grants from the NHLBI, RO1-HL-23140 and P50-HL-26215.

Three approaches taken in analyzing the data merit comment: (1) analysis of within-group and among-group variance in tabular data for baseline parameters in animals assigned to different intervention groups; (2) analysis of possible independence of the standard error of the mean and sample size in tabular data; and (3) analysis of variability between sequential radioactive microsphere regional myocardial flow measurements.

In several papers a striking lack of variability was observed among groups of dogs relative to the within-group variability. When possible independence of the standard error of the mean and sample size was examined, two papers were found to contain variables for which the standard error of the mean was virtually unchanged for groups with substantially different sample sizes. In regard to the measurement of regional myocardial blood flow, fractional differences between sequential measurements in non-ischemic segments in one paper were less than the measurement error of the microsphere technique reported in the literature. Dr. Darsee was responsible for data collection and analysis in each instance cited in these five papers. It is noteworthy that some of the dogs that allegedly received radioactive microsphere injections for determination of regional myocardial blood flow are not listed in the radioactivity log book. Furthermore, the few tissue blocks that are available from these dog experiments reveal no residual radioactivity.

An incident was noted in which a value in a published paper differed from that derived from the individual experimental data sheets in Dr. Darsee's laboratory notebook. (Appendix 5, Table 7). As a consequence, the magnitude of the reported reduction in infarct size in the paper was enhanced. In the preparation of summary tables in the notebook, values from individual experiments were altered or omitted.

b. Conclusions: The analyses summarized above call most of the tabular data in these five publications into question -- particularly when considered in the context of the data irregularities previously noted. Even if appropriate primary data were available for detailed review, it would remain difficult to vouch for the accuracy and validity of the papers in view of the extensive irregularities now apparent and the possibility of alterations in protocol which could not be appreciated by any form of data review. The primary responsibility for this situation rests with Dr. Darsee, but must be shared by his coauthors.

Dr. Darsee submitted to the Panel suggested explanations for the surprising statistical characteristics of his data. The Panel believes that the explanations offered by Dr. Darsee are unconvincing. Dr. Darsee also suggested to the Panel that many published papers from other laboratories share the statistical characteristics of his papers. Papers which were suggested by Dr. Darsee to be in this category were examined by Panel members and by statistical consultants; the Panel concludes that Dr. Darsee's suggestion is unfounded.

The fact that some of the dogs included in the published papers do not appear in the radioactive dog disposal log raises further doubts about those five papers; the absence of residual radioactivity in those few tissue blocks that are available confirms these suspicions.

2. Randomization of dogs

a. Findings: In the five papers referred to above, the Methods Sections imply that the dogs were assigned to control and experimental protocol: using prospective randomization procedures. (Appendix 5). In their March 25, 1982 letter to the Panel, Drs. Braunwald and Kloner acknowledge that there were numerous instances in which animals were used in more than one study. The dogs were shared among the five papers in such a way that prospective randomization could not have been used. The mean value for the necrosis/risk region for the control dogs used in a single paper was significantly less ($p<0.03$) than that for the control dogs used in multiple papers. The Methods Sections of papers using shared dogs do not represent all of the technical procedures in an accurate fashion. One paper implies that all animals were sacrificed at the end of a 6-hour reperfusion period; some of these animals were reported to have been used subsequently in other protocols. This would appear to be impossible.

b. Conclusions: These papers describe randomization procedures and some experimental methods in a manner that does not represent the actual procedures and methods in a completely accurate way.

Responsibility for the laboratory's use and reporting of randomization and experimental procedures ultimately resides with Drs. Kloner and Braunwald. In the studies of Dr. Darsee in which selection of retrospective control dogs was apparently not undertaken in a random manner, Dr. Darsee would appear to be accountable for any bias that might be present in the selection process.

F. Reporting of Regional Myocardial Blood Flow on a Heart that was Not Analyzed

1. Findings: On May 14, Dr. Darsee was observed doing a "Pacing Study" experiment (dog No. D-35). At the end of the day, the dog was dead, and it was observed by laboratory personnel to have been placed in a disposal bag without the heart having been removed. Because of suspicions that there were instances in which Dr. Darsee was reporting myocardial blood flow data on hearts that had not been analyzed, a member of the laboratory staff removed the heart and it was stored. In December 1981, the data book for this study was noted to contain output from a gamma counter and tapes for myocardial blood flow calculations labeled dog No. D-35. At this time (December 1981), Dr. Kloner was informed of the incident. The Panel examined this data book and found the gamma counter printout and flow calculations to be present and labeled dog No. D-35. The Panel was also shown the heart which had not been cut for determination of microsphere radioactivity. Dr. Kloner stated that he had discussed the incident with Dr. Braunwald in late 1981 but neither had informed the Harvard ad hoc Advisory Committee or the NHLBI Panel of it because they felt there was no proof that the excised heart was from dog No. D-35.

2. Conclusions: Although Dr. Darsee claims that he determined regional myocardial blood flow on biopsy samples (in contrast to the usual methods used in the laboratory), the Panel finds this explanation to be unlikely. The Panel does not believe that the blood flow data for dog D-35 in Dr. Darsee's notebook are valid since the heart was not removed from this animal. In addition, measurement of the heart for residual radioactivity has revealed levels indistinguishable from background levels. This incident, taken in conjunction with the other findings of the Panel, suggests further impropriety on the part of Dr. Darsee in the conduct of his experiments. It would have been preferable for Drs. Braunwald and Kloner to have informed Harvard's ad hoc Advisory Committee and this Panel of this incident, even though uncertainties about the identity of the stored heart and other aspects of the incident might have prevented absolute conclusions from being drawn at the time.

G. Notification of NIH of Non-Renewal of Dr. Darsee's Research Fellowship

1. Findings: On June 27, 1981, NIH was advised that Dr. Darsee would not be eligible for the third year of his Research Service Award. That advice was in the form of a letter signed by Dr. Darsee in which he asserted the following reasons for his action:

> "In the not-too-distant future I will be accepting a new position; both my academic responsibilities and my salary will change such that I will no longer be eligible for the NIH stipend... that now supports me."
>
> "I still have some writing to do in this laboratory and the precise date on which my responsibilities and source of income will change is flexible; however, I still wish to terminate the Research Service Award as of July 1, 1981 for the sake of simplicity."

Although this letter is written on official Brigham and Women's Hospital stationery, the extent to which Brigham and Women's Hospital officials were aware of the content of this letter is not clear to the Panel. On June 30, 1981, in response to a June 25 request from Dr. Braunwald, the Director, Office of Grants and Contracts, Brigham and Women's Hospital wrote the following letter to NIH:

> "Dr. John R. Darsee will be terminating his appointment at the Peter Bent Brigham Hospital, a Division of Brigham and Women's Hospital, effective July 1, 1981. Therefore, we are relinquishing all claims to year-03 of the referenced project."

2. Conclusions: The Panel believes that the letter of June 27, 1981 is, at the least, misleading and that the letter of June 30, 1981 is uninformative. Although Brigham and Women's Hospital officials complied with their legal obligations, and there were no official NIH guidelines indicating that they should have informed NIH of the basis for their action, the Panel believes that it would have been preferable nevertheless

for them to have done so. NIH has broad responsibilities to monitor its programs in an effective and prudent manner. By failing to notify NIH, Brigham and Women's Hospital impaired the ability of NIH to take appropriate remedial action.

H. Supervision and Instruction in the Laboratory

1. Findings: The procedures and practices in the Cardiac Research Laboratory - both prior to and after May 22, 1981 - have been examined. The laboratory has had an outstanding record of productivity since its establishment by Dr. Braunwald in 1973. It has provided research training and experience for many young investigators.

At about the same time that Dr. Darsee arrived at the laboratory in July 1979, responsibility for the day-to-day operation of the laboratory was transferred from Dr. Peter Maroko to Dr. Robert Kloner. Although the complement of postdoctoral research fellows remained relatively constant (approximately four at any one time), the nature of the training experience changed in some respects. Whereas during Dr. Maroko's last year in the laboratory all of the fellows participated in most research protocols, and thus shared many common experiences, the new procedure encouraged each fellow to develop and pursue his own studies. The formal orientation program to acquaint new fellows with laboratory procedures was conducted by Dr. Kloner and typically lasted four days. Continuing supervision and instruction occurred primarily on an informal basis, without regular didactic presentation of topics such as hypothesis testing, research design, randomization, adequacy of controls, data interpretation, and statistical techniques. Dr. Kloner reviewed tracings and other raw data with fellows individually within a day or two of the experiment; data meetings were held with Dr. Kloner every 2 weeks to discuss all ongoing studies and results; formal meetings to discuss studies and results with Dr. Braunwald were also held. The latter were scheduled to be held every 6-8 weeks but their regularity and frequency was variable in 1979-81. Drs. Kloner and Braunwald met twice weekly to discuss the results of individual experiments.

The review of data included in abstracts and papers to be submitted for publication did not routinely involve reexamination of individual records, but often relied on the ongoing examination of raw data and tracings conducted at an earlier time. Log books were not carefully maintained in the laboratory and were not reviewed by senior staff, the fellows' individual notebooks being regarded as the primary record. This practice has now changed and fellows are currently required to enter specific data for each experiment conducted in the general log book as well as in their personal records.

Raw data and tracings were not retained by Dr. Darsee and he stated he was never specifically instructed to do so. While it appears correct that no specific instructions were issued, the Panel's visits to the laboratory indicate that Dr. Darsee's contemporaries in the laboratory did understand that they were to retain and file their own records, and did so. No shortage of filing space existed at the laboratory.

In terms of newly adopted or reinforced laboratory policies, changes have been instituted over the last year with respect to: log books; adoption and implementation of new protocols; use of a standard laboratory manual; recording of data and retention of records; data review; and allocation of dog costs to specific grants or contracts.

2. Conclusions: Dr. Braunwald's many duties are complex, burdensome, and time-consuming. They involve a number of geographically discrete institutions. His accomplishments in discharging them are impressive and reflect his deep commitment, imaginative and productive approach, and remarkable efficiency, organization, and self-discipline. Nonetheless, these multiple responsibilities made it difficult for him to supervise the overall operation of the research laboratory in 1979-81. The clinical, teaching, and administrative duties which he has assumed do not allow sufficient time for direct participation in the research laboratory.

Dr. Kloner is a well-trained and capable scientist who was assigned major day-to-day responsibility for a busy research laboratory with several research fellows at an unusually early stage in his research career. He has made a genuine effort to discharge his duties in a careful manner but they appear to be too numerous and varied. He has not had an opportunity to assume these responsibilities in a gradual, phased manner.

In many respects supervision in the Cardiac Research Laboratory at the Brigham and Women's Hospital has been adequate and comparable to that provided in similar laboratories at other institutions. General supervisory practices have been based, appropriately, on an assumption of honesty and trust between supervisors and trainees. With the exception of Dr. Darsee, recent trainees have retained original records in customary and acceptable fashion. They have been allowed appropriately increasing independence as their laboratory experience has progressed. Because most incoming trainees have had superior records of previous accomplishment, they have perhaps been capable of progressing more rapidly than trainees in many other laboratories.

At the same time, the laboratory cannot be said to have performed optimally in ensuring appropriate handling, retention, and reporting of primary data after their collection by all fellows. A hurried pace and emphasis on productivity, coupled with limited interaction with senior scientists, has contributed to the disappointing events relating to randomization and reporting outlined in section VIII-E. Although a detailed review of papers published by the laboratory since 1979 in which Dr. Darsee was not a coauthor has not been conducted, the Panel was told of one published paper in which some controls from a previously published paper by the same investigator were used, with Dr. Kloner's approval, without the text so specifying.

The supervisory practices in the laboratory, while in no way responsible for Dr. Darsee's misconduct, may have contributed inadvertently to the ease with which he was able to produce fabricated data and to the subsequent difficulty in documenting the extent of the problem.

As a result of the unfortunate events reported since May 1981, a genuine and determined effort has been made to improve supervision in all phases of the laboratory's operation.

I. Report of the Ad Hoc Advisory Committee to the Dean of the Harvard Medical School on Dishonesty in Scientific Research

1. Findings: This committee was appointed by Dean Tosteson, with President Bok's concurrence, in November 1981 and consisted of five members of the Harvard faculty (three from the Medical School, one from the Faculty of Arts and Sciences, and one from the Law School) and three members from other institutions. Two meetings were held and a subcommittee visited the Cardiac Research Laboratory. The committee's final report was issued on January 25, 1982. Two important suggestions were made by this committee as to how the institutional response could have been improved. First, a small committee of senior professors from within the university, but outside the involved department, should have been consulted immediately after the discovery in May. Such a committee could have shared the burden with the dean, the chairman of the department, and the laboratory director, and offered objective advice concerning the management of the problem. Second, at the same time, a systematic search should have been conducted to identify all individuals within the institution with whom Dr. Darsee had collaborated, and these persons should have been informed confidentially of the allegations brought against him. The Panel concurs with these suggestions and believes that co-investigators outside the institution should also have been notified.

This Committee also suggested that laboratories, universities, publication boards, review bodies, and granting agencies should give consideration to the quality and significance of research as well as its quantity and visibility. The Panel also concurs with this suggestion.

This Panel's findings are at variance with the following conclusion of the Committee:

> "A systematic review of all work involving Dr. Darsee has been conducted by Drs. Kloner and Braunwald. Primary data have been reviewed and Dr. Darsee's co-workers have been questioned. The committee is convinced that Drs. Braunwald and Kloner have documented the extent of the irregularities and properly concluded that the previously published work from the laboratory in which Dr. Darsee was a member of the group is accurate. None of the work under suspicion has been published."

2. Conclusions: It would have been preferable for Harvard officials to have insisted on an exhaustive report without unrealistic time constraints. The limited time spent by the committee on their deliberations hampered an in-depth investigation that might have led to discovery of the full extent of data irregularities. It would also have been preferable for

the committee not to have relied so heavily on the internal audit conducted by Drs. Braunwald and Kloner. An audit by persons outside the involved laboratory and department could have led to an earlier resolution of these unfortunate events. In the Panel's view, the developments of late October and early November should also have led to the involvement of NHLBI.

J. Current Status of the AMPIM Study

1. Findings: In April 1982, a copy of the AMPIM protocol was sent to the laboratories at Johns Hopkins, Duke University, and Duke University/VA Hospital to learn how these laboratories conducting dog experiments interpreted certain aspects of the protocol. The protocol included the following:

> "Blinding of Post-Mortem Analysis - All tissue samples, blood flow data, and other data associated with the post-mortem analyses will be labeled with the animal ID number only. No indication of the experimental group to which the animal has been assigned will be given to the person who performs the post-mortem analyses."

As noted earlier in this report (VIII-D), a departure from the protocol occurred in all dog experiments at Brigham and Women's Hospital. Dr. Darsee carried out the surgery, randomized the dogs, and did all but one of the post-mortem analyses in an unblinded fashion. At Duke University, the same person was involved in the surgery, randomization of the dogs, and all of the post-mortem analyses. In this case, the group at Duke University states that labeling of the heart with only the ID number plus the analysis of several hearts at any one time effectively blinded the post-mortem analyses. At Johns Hopkins, separate persons were responsible for randomization of the dogs and the post-mortem analyses, and these analyses were carried out in a blinded manner. At Duke University/VA Hospital 3 professional staff members performed the post-mortem analyses and they were involved in the surgical and randomization procedures as well. However, tissue samples were appropriately labeled and several hearts examined at any one time so that, in effect, the post-mortem analyses were blinded.

2. Conclusions: The AMPIM protocol is interpreted by the Panel as implying that different persons should do the post-mortem analysis as compared to those doing the animal surgery and randomization. This procedure was not utilized at the Brigham and Women's Hospital, at Duke University, or at the Duke/VA Hospital. As a result, blinding of the study was not rigorous. The AMPIM Study was intended by the National Heart, Lung, and Blood Advisory Council to be a model study to demonstrate whether drugs could limit ischemic damage to the heart when the study was conducted concurrently in multiple centers in a rigorously blinded manner. Lack of rigorous blinding has resulted in a study that does not meet its original expectations. Thus, repetition of the portion of the study at Brigham and Women's Hospital would not accomplish the original purposes of the study.

K. Summary

1. Investigation of Data Irregularities:

In May 1981, Dr. John Darsee was involved in an episode of data fabrication in regard to a proposed abstract. This episode is acknowledged and not in dispute.

Disturbingly low variability has now been identified in all measurements performed by Dr. Darsee in the AMPIM Study before the acknowledged data fabrication in May 1981. Raw data needed to verify measurements of blood flow and hemodynamics have not been retained. Measurements of residual radioactivity in histologic tissue blocks indicate that radioactive microspheres were not injected for determination of regional myocardial blood flows prior to May 22, 1981 despite the fact that values for blood flows were reported. Reported large favorable effects of ibuprofen and verapamil on infarct size in the Brigham and Women's AMPIM studies are not deemed reliable.

Reported measurements of creatine kinase infarct size in another study also lack credibility.

There are statistical aberrations and apparent errors in tabular data provided by Dr. Darsee in several published papers from the Brigham and Women's laboratory in which he is a coauthor. These call into serious question both the reported data and the conclusions derived therefrom. The papers also include misleading descriptions of randomization procedures. The "methods" sections do not adequately describe all relevant procedures performed on each animal in each paper.

In December 1981, data purporting to represent radioactive microsphere measurements of regional blood flow were found in one of Dr. Darsee's laboratory notebooks for an animal which had been observed to be discarded with the heart in situ several months earlier. Also, many of Dr. Darsee's dogs which purportedly received radioactive microsphere injections do not appear in the laboratory's radioactive disposal log; measured levels of residual radioactivity in cardiac tissue from those animals are essentially the same as background levels.

Although he admits the episode of data fabrication in May 1981, Dr. Darsee denies any other wrongdoing. Dr. Darsee's opening statement to the Panel, his oral testimony, his written answers to submitted questions, and other written material that he submitted have not provided a satisfactory explanation for any of the numerous and varied events causing concern. The Panel's findings strongly reinforce and extend the conclusion of the ad hoc Advisory Committee to the Dean of the Harvard Medical School that Dr. Darsee "submitted to his superiors and to the scientific community unusual results which are highly suspect" and that he "was negligent in his responsibility to secure and maintain original data to substantiate his alleged results."

None of the Panel's findings suggest the involvement of any individual other than Dr. Darsee in the data fabrication at the Brigham and Women's laboratory.

2. Institutional Responses

When the acknowledged episode of data falsification by Dr. Darsee was uncovered in May 1981, Dr. Braunwald acted promptly to terminate Dr. Darsee's Harvard appointment and NIH fellowship. He appropriately informed administrative officials at Harvard and the Brigham and Women's Hospital of the circumstances of the incident. Some of the institution's responses at that point led to results that can now be appreciated as unfortunate. Investigation of the incident and auditing of Dr. Darsee's other work were entrusted by Dean Tosteson to Dr. Darsee's coauthors, Drs. Braunwald and Kloner, rather than to individuals from outside the department. Also, no systematic effort was made to identify and inform Dr. Darsee's other collaborators, including the AMPIM co-investigators, of the situation.

The internal audit conducted between May and October 1981 detected no evidence of questionable data or practices by Dr. Darsee, other than the single instance of data falsification he had admitted in May. Additional questionable data were identified by Drs. Kloner and Braunwald in October and November 1981 after they were alerted by NHLBI staff and AMPIM co-investigators to possible problems. The full extent of questionable data, including some in Dr. Darsee's published papers, was still not identified by the 1981 internal audit, however. It is apparent in retrospect that the internal audit conducted by Drs. Kloner and Braunwald was insufficiently rigorous and definitive. Major responsibility for this must be shared by Dean Tosteson; the Panel believes that it was unrealistic to expect coauthors of the work being scrutinized to be solely responsible for such an audit. It would have been preferable for the Dean to have had an early thorough, independent audit conducted by individuals who were not members of Dr. Braunwald's laboratory or department.

Following disclosure in October 1981 of other questionable data, Dean Tosteson established a special committee to investigate the entire situation. This committee identified no additional instances of questionable data and, in effect, ratified the results of the internal audit. The Panel believes that it would have been appropriate for a committee formed at this point to have involved the NHLBI. Above all, it would have been preferable for Harvard to have ensured that the committee had sufficient time and opportunity to conduct a thorough and exhaustive evaluation.

In December 1981, NHLBI established its own Panel to investigate the situation. At all times, Drs. Tosteson, Braunwald, Kloner, and all other officials and employees at Harvard and the Brigham and Women's Hospital cooperated fully and promptly in providing all requested information and in facilitating the Panel's study in every way possible.

3. Supervision of Procedures in the Laboratory

Apart from the professional misconduct of Dr. John Darsee, the integrity and scientific capabilities of the professional and technical staff of the cardiovascular laboratory at the Brigham and Women's Hospital are of a high order. In reaching its conclusions concerning the actions of Dr. Braunwald,

Dean Tosteson, and others involved in the unfolding of events surrounding Dr. Darsee's misconduct, the Panel members recognize that they have had the benefit of viewing these events from a later and much more informed vantage point. Also, as a consequence of Dr. Darsee's misconduct, this laboratory has undergone an unusually rigorous scrutiny. The Panel is of the opinion that the circumstances prevailing throughout Dr. Darsee's period in the laboratory, while not responsible for or in any way condoning his misdeeds, helped to create an environment that may have inhibited their being uncovered. In particular, mention may be made of the competing obligations of Dr. Braunwald, Dr. Kloner's lack of previous experience in supervising a busy laboratory, and a hurried laboratory pace. As a consequence, some laboratory procedures were less than optimal, in regard to items such as ensuring proper analysis and verification of raw data, appropriateness of controls and randomization techniques, and attention to accurate description of methodology in manuscripts.

Since these unfortunate events, appropriate changes have been implemented by Drs. Braunwald and Kloner to improve the overall operation of the laboratory.

IX. Recommendations

General - Fraud in research, as in every aspect of human affairs, has been documented over many years, and debate may take place on the particular circumstances of each occurrence and the motivation or psychological aberration of the individual concerned. Such incidents continue to emphasize the obligations of all concerned in research for its proper conduct, including supervision of trainees, appropriate retention of primary data, regular scrutiny by experienced scientists of primary data and subsequent calculations, and encouragement of free scientific interchange and exchange of data. Each author named on a manuscript should be prepared to accept responsibility for its content. If this is not possible, the limits of the individual's responsibility should be defined in a footnote, or the individual should be recognized in the acknowledgments rather than as a coauthor. The Panel concludes, however, that even under optimal conditions, the possibility will still exist for data falsification as has occurred in the circumstances involving Dr. John Darsee. For this reason formulation of a strict set of rigid rules would be of little value, and could have a serious negative effect on creativity in science.

When data falsification occurs or is alleged to have occurred, a number of actions should be taken, and the Panel endorses the following guidelines as representative of an appropriate and desirable approach. If not already done, the circumstances should be brought to the attention of the principal investigator in the laboratory and the chairman of the administrative unit. The latter should inform the dean or comparable administrative official, who in turn should appoint immediately a small committee of knowledgeable individuals from within the institution, but outside the involved administrative unit. This committee should investigate the allegations promptly and determine if there are substantive grounds for suspecting serious irregularities. This committee should reach one of three conclusions in a timely manner. First, that the allegation is unfounded; second, that the allegation is proven and should be dealt with appropriately; or third, that a more deliberative and exhaustive investigation is needed by a special panel with expertise in the area of dispute. Depending on specific circumstances, including the seniority and position of the individuals involved and the complexity of the scientific questions at issue, members from outside the institution might be included at this stage of the investigation.

When it is clear that serious wrongdoing has occurred, the other scientific activities of the individual should be investigated and the following should be informed:

1. Collaborators in ongoing research projects.

2. Coauthors of papers whether submitted for publication, in press, or published.

3. Granting agencies with which the individual was involved at the time the data falsification occurred and granting agencies funding the individual's current research.

Each of these individuals and granting agencies must use proper moral and legal judgment in the way they respond to such information.

If the complexity of the case leads to a conclusion that the issue cannot be resolved rapidly, the responsible institutional official should decide whether collaborators, coauthors, and/or granting agencies should be notified in confidence of the situation at this time. The Panel recognizes that the diversity of incidents and circumstances may require a variety of responses. The investigative panel should conduct a comprehensive evaluation to decide whether or not the allegations are proven and determine the full extent of any misdeeds. The final conclusions of the panel should be communicated promptly to any previously notified collaborators, coauthors, and granting agencies. If wrongdoing is found to have occurred, all other relevant individuals and agencies should be informed as well. The investigations must be deliberate and thorough at every stage of the process. The speed of the investigation should not be allowed to compromise its thoroughness.

Specific Recommendations:

1. In regard to Dr. Darsee, the Panel recommends that the NIH proceed with debarment for the maximum allowable time. It is noted that this procedure involves an opportunity for a formal hearing should Dr. Darsee decide to appeal. In addition, given the apparent fabrication by Dr. Darsee of research data submitted to the Federal government, the Panel recommends that the Department of Health and Human Services consider referring this matter to the Department of Justice.

2. The Panel recommends that Dr. Darsee's request to repeat the Brigham and Women's portion of the AMPIM study be denied.

3. The Panel recommends that NIH officials, in considering present and future support to the Cardiac Research Laboratory at the Brigham and Women's Hospital, confirm the Panel's impression that current laboratory procedures and supervision are adequate.

4. The Panel recommends that the NHLBI not approve the request of the Brigham and Women's Hospital to repeat their portion of the AMPIM study.

5. The Panel recommends that the results obtained by Duke University, Johns Hopkins, Duke University/VA Hospital and the St. Louis University in the AMPIM study be published at the discretion of each participating laboratory. This procedure will allow each group of investigators to describe in full the protocol that was followed and the results obtained.

6. The Panel recommends that the NIH should develop and promulgate guidelines for reporting scientific misconduct to the appropriate bureau, institute or division. These guidelines should describe the degree of confidentiality that can be ensured.

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Memorandum

Date JAN 5 1990

From Associate Director for Extramural Research and Training

Subject Investigation of Alleged Misconduct at Brigham and Women's Hospital and Proposed Actions--DECISION

To Director, NIH

SUMMARY

An investigation by the National Heart, Lung, and Blood Institute (NHLBI) has confirmed that research data were fabricated in the course of an NHLBI-funded contract to the Cardiac Research Laboratory at the Brigham and Women's Hospital, Harvard Medical School. The contract was one of five awards funding a cooperative study entitled "Animal Models for Use in the Assessment of Therapy to Protect Ischemic Myocardium" (AMPIM). The data falsifications were judged to be acts of a single individual, Dr. John R. Darsee, who was supported by an individual National Research Service Award (post-doctoral fellowship) from the NHLBI.

A panel of senior non-Federal scientists, convened by the NHLBI, investigated the allegations of data falsification. Following receipt of the Panel's report, the National Institutes of Health (NIH) received additional information suggesting that the falsifications occurring in the course of the contract-funded research were not isolated events but part of a pattern involving Dr. Darsee's earlier research at Emory University. Thus a final assessment of the full significance of Dr. Darsee's actions cannot be made until the additional evidence is weighed and a second investigation, as recommended below, is completed. However, the findings at Harvard with respect to Dr. Darsee's actions are sufficiently serious to warrant the initiation of a debarment proceeding for Dr. Darsee.

The Panel's recommendations regarding the incidents at Harvard are discussed in the body of this memorandum. The recommendations of a committee of senior NIH staff, which reviewed both the report and the comments of all principals, are as follows:

(1) That you ask the Secretary to initiate proceedings to debar John R. Darsee, M.D. from both financial assistance (grants and cooperative agreements) and contract funding for a period of ten years.

(2) That Dr. Darsee be excluded from service on any NIH peer review or program advisory committee for a period of ten years from the date of this memorandum or until the end of any period of debarment, whichever is longer; and that any NIH component proposing him for subsequent appointment provide the Director, NIH, with a detailed written justification.

(3) That NIH conduct a thorough review of Dr. Darsee's research and publications at Emory University.

(4) That NHLBI conduct an on-site review of supervisory practices and research procedures of the Cardiac Research Laboratory at Brigham and Women's Hospital in approximately one year's time to confirm the Panel's impression that current laboratory procedures and supervision are adequate.

(5) That the request of Dr. Darsee to repeat the study be denied, and that the request of the Brigham and Women's Hospital to repeat its portion of the AMPIM study be denied.

(6) That the NHLBI recover the funds awarded to the Brigham and Women's Hospital for its portion of the AMPIM study in the amount of $122,371.

(7) That the results obtained by Duke University, Johns Hopkins, Duke University/VA Hospital and the St. Louis University in the AMPIM study be published at the discretion of each participating laboratory.

BACKGROUND

In October 1981 the staff of the National Heart, Lung, and Blood Institute became aware of problems involving the comparability of data reported by investigators in a cooperative study entitled "Animal Models for Use in the Assessment of Therapy to Protect Ischemic Myocardium." This study, referred to by the acronym AMPIM, was begun in Fiscal Year 1979 with the award of five contracts.

Late in October, it became apparent that some of the data submitted by the investigators at the Brigham and Women's Hospital were inconsistent with those submitted by other centers and were not consonant with results obtained in several prior studies. The data in question at that time were measurements of regional myocardial blood flows and heart weights. The former showed a remarkable lack of variability and the latter a discordant relationship of left ventricular weight to dog weight. At a meeting of the AMPIM investigators with NHLBI staff on October 30 Dr. Robert Kloner, an investigator at the Brigham laboratory, divulged that a postdoctoral research fellow who had a major role in the

AMPIM studies in this laboratory had been found about five months earlier to have fabricated experimental records and research data for a project not directly related to the AMPIM contract. In a letter of November 3 Drs. Braunwald and Kloner indicated that they could not vouch or take responsibility for the data provided by this individual.

The episode of observed and admitted data falsification occurred in May 1981. The individual who committed these acts was John R. Darsee, M.D., who was supported by a fellowship award from the NHLBI (F32-HL5958). The sponsor was Dr. Eugene Braunwald. The award commenced on July 1, 1979, and was to have a three-year duration. However, at the request of the institution, the award was not renewed after June 30, 1981; there was no indication given to NHLBI that this was anything other than a routine action.

Although the senior staff in the Cardiac Research Laboratory at the Brigham and Women's Hospital (Drs. Braunwald and Kloner) had been conducting their own audit of these events, and although the Dean of the Harvard Medical School had on November 25, 1981, established an external review committee, the NHLBI believed that an expert panel commissioned by the granting agency itself was warranted. Such a panel was constituted in December.

The individuals appointed to the Panel were:

Dr. Howard E. Morgan
Evan Pugh Professor of Physiology
Department of Physiology
The Pennsylvania State University
College of Medicine
Hershey, Pennsylvania 17033

Dr. Francis J. Klocke
Chief
Division of Cardiology
Professor of Medicine and Physiology
State University of New York
at Buffalo
Buffalo, New York 14215

Dr. John T. Shepherd
Director for Education
Mayo Foundation
Dean
Mayo Medical School
Rochester, Minnesota 55905

Dr. C. Kern Wildenthal
Dean
Southwestern Medical School
The University of Texas Health
Science Center at Dallas
Dallas, Texas 75235

Dr. Morgan served as Chairman of the Panel and Dr. Jerome G. Green, Director of the Division of Extramural Affairs at the NHLBI, served as Executive Secretary.

Staff members who worked with or observed the Panel were:

Dr. Kent R. Bailey, NHLBI
Dr. David L. DeMets, NHLBI
Mr. Donald C. Gall, Division of Management Survey and Review, NIH
Mr. Robert B. Lanman, Office of the General Counsel
Mr. Joel M. Mangel, Office of the General Counsel
Dr. Eugene R. Passamani, NHLBI
Mr. Richard J. Riseberg, Office of the General Counsel
Dr. Richard P. Schwarz, NHLBI
Mr. Patrick M. Sullivan, NHLBI

In the letter of invitation to the four members of the Panel the Institute's charge to them was outlined, as follows:

"The NHLBI Panel is to review circumstances and events in the Cardiac Research Laboratory at the Brigham and Women's Hospital relating to alleged scientific misconduct.

"The Panel is asked to:

describe the nature, extent, and seriousness of the alleged misconduct

identify the NIH support instruments and awards that were involved

describe and assess the relevant practices, procedures, and supervision that existed, and that now exists, in these laboratories

review the reporting and accountability procedures and responsibilities that exist and that were practiced with respect to Brigham and Women's Hospital, Harvard Medical School and the NHLBI

recommend any additional steps and precautions that seem indicated and prudent to deal with the present circumstances or future occurrences and that might serve to prevent such incidents."

The members of the panel were asked to assess overall supervision and instruction of research fellows in the Cardiac Research Laboratory at the Brigham and Women's Hospital during the period 1979, when Dr. Darsee began work at Harvard, to the present. The Panel was not asked to investigate Dr. Darsee's scientific activities prior to his arrival in Boston. Detailed review of published papers from the laboratory was restricted to those for which Dr. Darsee was an author.

The methods and procedures employed by the Panel are described in its Report, which also describes the purpose and design of the AMPIM study, the chronology of events, and the findings, conclusions, and recommendations of the Panel.

The Panel's investigation, conducted between February and June, 1982 provided ample opportunity for all parties involved to present oral and written information regarding the events in question. Upon receiving the Panel's report I convened a staff committee to review it and determine what the next steps should be. At a meeting held on July 30, Dr. Jerome Green analyzed the major issues in dispute between the Panel and principals involved. Senior management officials participating in that discussion included:

Claude Lenfant, M.D., Director, National Heart, Lung, and Blood Institute

Betty Pickett, Ph.D., then Deputy Director, National Institute of Child Health and Human Development, now Director, Division of Research Resources

James F. O'Donnell, Ph.D., Acting Director, Division of Research Resources

Mr. Carl Fretts, Director, Division of Contracts and Grants

Mr. Howard Hyatt, Director, Division of Management Survey and Review

Messrs. Richard Riseberg and Robert Lanman of the Office of the General Counsel also participated, as did other staff from NHLBI, DMSR, and the Office of Extramural Research and Training

The group concluded that it would be useful to have the comments of all principals on the final report. In particular, it was noted that Dr. Braunwald had reiterated his earlier request to meet personally with those who would be recommending or taking action on the Panel's report. It was the consensus of those present that it would be preferable to request written comments from all involved parties. Accordingly, I wrote to Dr. Kloner, Dr. Braunwald, Dean Tosteson, and Dr. Darsee's attorney requesting their comments.

When those comments were received I convened a second meeting on October 5, involving the persons attending the previous meeting as well as Darrel A. Regier, M.D., Director, Division of Biometry and Epidemiology, NIMH, and Mr. Storm Whaley, Associate Director for Communications, NIH. The highlights of that discussion are summarized below.

DISCUSSION

The first topic of discussion was a letter from Dr. Darsee stating that, although he had no recollection of falsifying any research data, he acknowledged that the Panel's inquiry had established both the fact of falsification and his personal role. In view of the highly personal nature of that letter, and Dr. Darsee's request that it not be

disclosed, I subsequently requested that he prepare a public statement to replace the original letter. Although Dr. Darsee's attorney indicated to a member of my staff that such a letter would probably be provided, as of this writing I have received no further communication from Dr. Darsee. It should be noted that Dr. Darsee's letter was a significant departure from his earlier assertions to the Panel that he had not engaged in any deceptive or irregular practices other than the single incident in May 1981.

The Panel's recommendations with regard to Dr. Darsee were as follows:

"1. In regard to Dr. Darsee, the Panel recommends that the NIH proceed with debarment for the maximum allowable time. It is noted that this procedure involves an opportunity for a formal hearing should Dr. Darsee decide to appeal. In addition, given the apparent fabrication by Dr. Darsee of research data submitted to the Federal government, the Panel recommends that the Department of Health and Human Services (DHHS) consider referring this matter to the Department of Justice.

"2. The Panel recommends that Dr. Darsee's request to repeat the Brigham and Women's portion of the AMPIM study be denied."

All those present at the meeting of the staff committee agreed that debarment is warranted, and they recommended that you request the Secretary to initiate proceedings to consider debarment of Dr. Darsee from DHHS financial assistance and contract funding for a period of seven to ten years. (The regulations do not specify a maximum or minimum period.) The suggested range reflects two schools of thought; the committee did not arrive at a consensus. Proponents of a lower-period of time argued that Dr. Darsee's actions render his unfit to engage in DHHS-sponsored research for the foreseeable future. Others felt that a ten-year or longer debarment would be excessively harsh in view of the fact that Dr. Darsee has completed only the early stages of a research career and might eventually prove capable of performing responsibly in research. It was noted that the debarment proceedings provide an opportunity for the Secretary to consider a variety of factors, and also that a debarred individual may apply to the Secretary to have the debarment lifted if circumstances warrant. I have considered these arguments at length and have reached the conclusion that the ten-year period is appropriate.

It should be noted that although the Panel recommended that DHHS consider referring this matter to the Department of Justice (DoJ), such referral is not optional. Any finding of possible criminal violation is routinely referred to the Justice Department by the HHS Office of the Inspector General (OIG). In fact, by memorandum dated August 5, the Division of Management Survey and Review (DMSR) has already transmitted to the OIG a copy of the Panel's report on the results of its investigation.

The Panel also recommended that Dr. Darsee not be permitted to repeat the study, nor should the Brigham and Women's Hospital be allowed to repeat its portion of the AMPIM study, as had been requested. The Panel stated, and the group agreed, that the study had not met its original expectations, and that repetition of that portion would not accomplish the original purposes of the study.

An additional sanction recommended by the staff committee, but not discussed in the Panel's report, would prohibit Dr. Darsee from serving on any NIH peer review or program advisory committee for a period of ten years from the date of this memorandum or until the end of any period of debarment, whichever is longer. In addition, any NIH component submitting a subsequent nomination for Dr. Darsee after this period should be required to provide the Director, NIH, with a written justification.

The staff committee also considered the question, raised in Dr. Braunwald's comments, of whether Dr. Darsee's earlier work at Emory University should be reviewed by the NIH. In view of the serious penalties already recommended, the group initially believed that such a review was not necessary to determine a course of action with regard to Dr. Darsee. We continue to believe that there is ample basis to initiate debarment proceedings now. However, Dr. Braunwald recently provided me with detailed analyses of research papers published by Dr. Darsee while he was at Emory. These analyses suggest that numerous fabrications may have occurred in Dr. Darsee's earlier research as well. In view of these developments, and the fact that Emory is conducting its own review, I shared that information with officials of NHLBI and other concerned staff including the Office for Protection from Research Risks (OPRR) noting that several of the projects at Emory involved human subjects. The consensus now is that the alleged misconduct at Emory is very serious and that NIH should undertake an investigation at Emory as soon as resources permit. This second investigation would focus primarily on the protection of human subjects, the veracity of published papers, and the disposition of Federal funds, although the possibility of further sanctions against Dr. Darsee cannot be ruled out.

The responses of Drs. Braunwald and Kloner present complex issues, for these individuals disagree with the Panel's report both on matters of fact and interpretation. In reviewing those comments, the staff committee affirmed that the Panel's interpretation of its charge was both correct and executed with commendable thoroughness. Dr. Lenfant, however, expressed strong reservations about the propriety of having directed the Panel to assess laboratory practices and supervision.

The Panel made several recommendations specifically concerning the work of Dr. Braunwald and Dr. Kloner. These are as follows:

"3. The Panel recommends that NIH officials, in considering present and future support to the Cardiac Research Laboratory at the Brigham and Women's Hospital, confirm the Panel's impression that current laboratory procedures and supervision are adequate.

"4. The Panel recommends that the NHLBI not approve the request of the Brigham and Women's Hospital to repeat their portion of the AMPIM study."

We advocate qualified acceptance of recommendation 3. The committee was persuaded that the measures described in the Panel's Report and in the responses of Drs. Braunwald and Kloner are reasonable and sufficient. We also recognize that no system of procedures and supervisory controls can offer absolute protection against willful deceit. It seems clear that an additional review of the Cardiac Research Laboratory in the immediate future would add little to the information already available, nor does there appear to be any basis for assuming that special reviews are needed for the indefinite future. Therefore, at its October 5 meeting the committee recommended that the NHLBI conduct a site visit one year hence to confirm that the procedures and supervision in the Cardiac Research Laboratory are appropriate. If that review confirms the Panel's opinion, no further special reviews should be undertaken.

The Panel's report in several places contains critical observations and comments about laboratory practices and supervision. Drs. Braunwald and Kloner have expressed strong objections and furnished detailed rebuttals. The staff committee recognized that the Panel's findings in this regard are matters of professional judgment and that equally qualified observers might form different impressions from the same information. Accordingly, the staff committee concluded that an attempt to make point-by-point acceptance or rejection of each specific finding or conclusion of the Panel in this area would not be warranted. Rather, we recommend that the comments of Drs. Braunwald and Kloner be appended in full to the Panel's Report, thereby ensuring that the public record associated with NIH decisions includes both perspectives.

As noted above, the NHLBI has concluded that no useful purpose would be served by repeating the Harvard portion of the AMPIM study. Accordingly, we concur with the recommendation that NHLBI not approve the request of the Brigham and Women's Hospital to repeat its portion of the AMPIM study. In addition, we recommend that NHLBI recover the $122,371 awarded for conduct of that portion of the study.

The Panel made the following recommendation with regard to results obtained by other laboratories collaborating in the AMPIM study:

> "5. The Panel recommends that the results obtained by Duke University, Johns Hopkins, Duke University/VA Hospital and the St. Louis University in the AMPIM study be published at the discretion of each participating laboratory. This procedure will allow each group of investigators to describe in full the protocol that was followed and the results obtained."

The committee concurs with this recommendation.

The committee was pleased to learn that Drs. Braunwald and Kloner have taken steps to retract papers they co-authored with Dr. Darsee involving data obtained by Dr. Darsee in the Cardiac Research Laboratory.

The committee reviewed two communications from Dean Tosteson. In a letter of September 17 he reported that an internal review complemented and confirmed earlier studies that concluded that Dr. Darsee had failed to perform certain procedures in the AMPIM study. Dean Tosteson's September 24 letter commented specifically on the NHLBI report and possible NIH actions. Dean Tosteson's letter was supportive of the request of Drs. Braunwald and Kloner to repeat their portion of the AMPIM study if the results would be useful, but conceded that may not be the case. He indicated his institution's willingness to enter into discussions regarding possible recovery of funds. He commented that Harvard would be troubled by the perception that further review of the Cardiac Research Laboratory is needed but would not object to special reviews. Finally, he described efforts under way at Harvard to deal with allegations or evidence of misconduct in science.

The Panel's sixth recommendation was discussed but not acted upon by the staff committee. It reads as follows:

> "6. The Panel recommends that the NIH should develop and promulgate guidelines for reporting scientific misconduct to the appropriate bureau, institute or division. These guidelines should describe the degree of confidentiality that can be ensured."

As you know, development of guidance for agency staff and awardees is a matter of high priority for the PHS and the subject of considerable effort on the part of my staff. The academic community is also addressing the need for more specific guidelines and procedures. Drs. Braunwald and Kloner are correct in observing that it is easier to critize in retrospect than to choose the appropriate course of action in advance. Their desire to be fair to Dr. Darsee and to ensure due process is commendable, but the fact remains that a large and costly study of great importance for a major public health problem was irrevocably compromised because of the failure to inform NHLBI and the AMPIM co-investigators.

Recommendations

1. That you request the Secretary to initiate proceedings to debar John R. Darsee, M.D. from eligibility for discretionary financial assistance and contract funding from DHHS for a period of ten years.

   Concur [signature] Non-concur_____ Date 1-5-85

2. That Dr. Darsee be excluded from service on any NIH peer review or program advisory committee for a period of ten years from the date of this memorandum or until the end of any period of debarment, whichever is longer; and further, that any NIH component submitting a subsequent nomination for Dr. Darsee to serve on such a committee be required to provide the Director, NIH, with a detailed written justification.

   Concur [signature] Non-concur_____ Date 1-5-85

3. That the OPRR and appropriate awarding units examine Dr. Darsee's research activities at Emory University including an assessment of whether the rights of human subjects or the terms of Emory's general assurance were violated.

   Concur [signature] Non-concur_____ Date 1-5-85

4. That NHLBI conduct an on-site review of supervisory practices and research procedures of the Cardiac Research Laboratory at the Brigham and Women's Hospital in approximately one year's time to confirm that the laboratory procedures and supervision are appropriate.

   Concur [signature] Non-concur_____ Date 1-5-85

5. That the requests of Dr. Darsee, and of the Brigham and Women's Hospital, to repeat their work on the AMPIM study be denied.

   Concur [signature] Non-concur_____ Date 1-5-85

6. That the NHLBI recover from Harvard Medical School the funds awarded for the Brigham and Women's Hospital portion of the AMPIM study, in the amount of $122,371.

   Concur [signature] Non-Concur_____ Date 1-5-85

Page 11 - Director, NIH

7. That the results obtained by Duke University, Johns Hopkins, Duke University/VA Hospital and the St. Louis University in the AMPIM study be published at the discretion of each participating laboratory.

Concur [illegible] Non-Concur_____ Date 1-5-83

William F. Raub

William F. Raub, Ph.D.

HHS NEWS

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOR RELEASE
February 16, 1983

NATIONAL INSTITUTES OF HEALTH
Dr. William Raub (301) 496-1096
Mary L. Miers (301) 496-5366

Assistant Secretary for Health Edward N. Brandt Jr., M.D., today announced initiation of proceedings to debar John R. Darsee, M.D., from eligibility for funding from the Department of Health and Human Services.

James B. Wyngaarden, M.D., director of the National Institutes of Health, had recommended these actions against Dr. Darsee, a former research fellow in the cardiac research laboratory at the Brigham and Women's Hospital, Harvard Medical School, in Boston.

The National Heart, Lung, and Blood Institute, an NIH component, has found that Dr. Darsee fabricated research findings while working on a study funded by an NHLBI contract. The study, called "Animal Models for Use in the Assessment of Therapy to Protect Ischemic Myocardium (AMPIM)," was designed to develop animal models to assess the efficacy of drugs to limit infarct (muscle damage) size in heart attack victims.

The investigation of Dr. Darsee's work was conducted by the NHBLI, assisted by a panel of distinguished academic cardiologists chaired by Howard E. Morgan, M.D., of the Pennsylvania State University College of Medicine.

In addition, NIH has decided to exclude Dr. Darsee from service on any peer review or program advisory committee for 10 years. After that time, any NIH component submitting a nomination for Dr. Darsee to serve on a committee must provide the NIH director with a detailed justification for the nomination.

(more)

NIH has denied Dr. Darsee's request that he repeat his portion of the ATPIM study. Although the Brigham and Women's Hospital offered to repeat its portion of the ATPIM study at no cost to the government, Dr. Wyngaarden has accepted the NHLBI panel's finding that this would not accomplish the study's purposes, and the Hospital will be asked to return the $122,371 awarded for its portion of the study.

In addition to allegations of misconduct on the part of Dr. Darsee, the NHLBI panel assessed relevant practices, procedures, and supervision in the Hospital's cardiac research laboratory. The panel found that some aspects of record keeping could have been improved, but noted that corrective steps have been taken.

NIH will also review Dr. Darsee's NIH-funded research at the Emory University School of Medicine, Atlanta, following receipt of a report from Emory, which is conducting its own investigation.

###

COMMENTS ON NIH REPORT

BRIGHAM AND WOMEN'S HOSPITAL - HARVARD MEDICAL SCHOOL - BETH ISRAEL HOSPITAL

EUGENE BRAUNWALD, M.D.
Hersey Professor of the Theory and Practice of Medicine

Chairman, Departments of Medicine
Physician-in-Chief
Brigham and Beth Israel Hospitals

MEMORANDUM

To: Members of the Committee of Senior NIH Staff Convened by Dr. Raub to Make Recommendations to Dr. Wyngaarden Concerning the Report of the National Heart, Lung and Blood Institute Special Panel to Review Alleged Misconduct at Brigham and Women's Hospital/Harvard Medical School, Boston, Massachusetts

From: Eugene Braunwald, M.D.

Subject: Comments Concerning NHLBI Panel's Report and Range of Possible NIH Actions

Date: September 21, 1982

I greatly appreciate the attention devoted by members of the NHLBI Special Panel and the NHLBI Staff to their study and I welcome the request of the NIH for comments on the Panel's Report. In reviewing the events of Dr. Darsee's misconduct at the Brigham and Women's Hospital/Harvard Medical School, the investigation of the Panel provides valuable groundwork for the construction of needed guidelines for institutional prevention of and response to scientific misconduct in the future. I believe deeply that the final actions taken by the NIH on this particular matter should serve to clear away the cloud over the Cardiac Research Laboratory at Brigham and Women's Hospital and its investigators, research fellows and technical staff who were not involved in any way, as they are the most clearly identifiable victims of a grave and extraordinary breach of trust by a single individual.

This Memorandum has six appendices attached. Appendix E sets out the changes and additions that I believe are necessary to make the chronology contained in Section VI-B of the Panel's Report an accurate statement of the

Reply to: Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115 (617) 732-6340

events pertinent to the Panel's inquiry. Appendix F contains a summary of my responses to the specific conclusions of the Report following the same format and organization as the Panel's Report, with references to this Memorandum wherever possible to avoid repetition.

CONTENTS

		Page
I.	General Comments	5
	A. The integrity and scientific capabilities of the professional and technical staff of the Brigham and Women's Cardiac Research Laboratory are of a high order.	5
	B. Faced with a lack of NIH guidelines, Brigham and Women's personnel complied with all legal obligations in investigating and reporting Dr. Darsee's misconduct.	5
	C. Brigham and Women's personnel have taken the initiative in investigating Dr. Darsee's misconduct.	7
	D. Supervision and operation of the Cardiac Research Laboratory are adequate.	9
II.	Specific Comments	10
	A. The charge to the Panel (or the Panel's interpretation of its charge) inappropriately limited investigation to the Cardiac Research Laboratory.	10
	B. The Report's suggestion that our investigation of Dr. Darsee's work was insufficiently rigorous ignores the extraordinary difficulty, illustrated by the Panel's own experience in this matter, of investigating possibly fraudulent research.	12
	C. Suggestions in the Report that the quality or quantity of supervision in the laboratory enhanced Dr. Darsee's ability to cheat or avoid detection are unjustified by the Panel's limited investigation.	15
	D. The Report does not state the complex considerations we faced in May 1981 which led us to make the decisions we did.	27
	E. The Panel's Report does not, but most certainly should, state that it was Dr. Kloner's finding that led us to conclude in October 1981 that Dr. Darsee's AMPIM data were not reliable.	33
	F. It is our hope that the laboratory will be authorized to repeat its portion of the AMPIM study unless, because of the lack of rigorous blinding at three of the four institutions, the AMPIM results are not to be used.	35

Page

G. Consistent with their obligations to the scientific community and as a precautionary measure, Drs. Braunwald and Kloner have requested the editors of journals which have published articles contributed by Dr. Darsee from Harvard to publish retractions of articles or sections of articles which are suspect. 35

Appendix A. Review of papers published by Dr. Darsee at Emory University

Appendix B. American Heart Association James B. Herrick Award to Dr. Braunwald

Appendix C. Letters of recommendation for Dr. Darsee

Appendix D. Letters to journals from Drs. Braunwald and Kloner

Appendix E. Requested Corrections and Additions to Section of the Report on the Chronology of Events at the Brigham and Women's Hospital

Appendix F. Response to the Findings and Conclusions Section of the Report

I. General Comments

There are several important conclusions indicated by the Report with which I am in agreement, some of which were also confirmed by the Ad Hoc Advisory Committee to the Dean of the Harvard Medical School on Dishonesty in Scientific Research which investigated this matter. Several of these conclusions are summarized below in order to provide an appropriate context for my discussion of other comments in the Panel's Report with which I cannot wholly concur.

A. The integrity and scientific capabilities of the professional and technical staff of the Brigham and Women's Cardiac Research Laboratory are of a high order.

The Report concludes that "[a]part from the professional misconduct of Dr. John Darsee, the integrity and scientific capabilities of the professional and technical staff of the Cardiovascular Laboratory at the Brigham and Women's Hospital are of a high order" (page 32), and that "[n]one of the Panel's findings suggest the involvement of any individual other than Dr. Darsee in the data fabrication at the Brigham and Women's Laboratory" (page 31).

B. Faced with a lack of NIH guidelines, Brigham and Women's personnel complied with all legal obligations in investigating and reporting Dr. Darsee's misconduct.

The Report, although questioning in retrospect the advisability of some aspects of the Hospital's response to its discovery of a single, admitted incident of data fabrication, confirms that ". . . Hospital officials complied with their legal obligations, and there were no official NIH guidelines indicating that they should have informed NIH" of their reasons for terminating the fellowship of Dr. Darsee. (Page 26.) Similarly, the Report acknowledges that " . . . NIH has not issued guidelines concerning the responsibility of grantees, sponsors, or contractors to report actual or alleged misconduct . . ." (Page 19.)

The complexity of the issues faced by us in balancing the interests involved and responding to the circumstances as they gradually became known in the absence of such guidelines is reflected in the testimony of Dr. Donald Fredrickson, former Director of the NIH, and the late Dr. Philip Handler, former President of the National Academy of Sciences, before the Congressional Subcommittee on Investigations on Oversight on March 31 and April 1, 1981, less than two months before we discovered Dr. Darsee's misconduct:

> Dr. Fredrickson: "To be sure also because human beings are involved; ambitious human beings, seeking honor and prestige and we can easily injure them; we can destroy them for a whole career. We can cast them out of science and one has to be extraordinarily careful in the exercise of judgment and the kind of justice that is harsh and exists in dealing with these matters."
>
> Mr. Gore: "Did you wish to comment, Dr. Handler?"
>
> Dr. Handler: "I just couldn't agree more with what Dr. Fredrickson has just said. But I will admit with you, sir, the absence of any sense of what due process should be when some suspicion is aroused. We have never adopted standardized procedures of any kind to deal with these isolated events. We have no courts, no sets of courts, no understandings among ourselves as to how any one such incident shall be treated.
>
> "We have left them, one at a time, as they have occurred, to the best judgment of the institutions within which such events have transpired when one thought they had. And, in the end, word of such misdeeds which attaches to the name of the individual invariably destroys him in the career on which he had embarked. We think that suffices; more than suffices."

Later in the same hearing:

> Dr. Fredrickson: "Nothing like that [the discrediting of a scientist involved in data falsification] moves immediately. I don't think you get that speed. It moves very deliberately, because one must be sure. If the result is the exclusion of that individual from science, that is the ultimate penalty, and one must proceed carefully, because there are ambiguities, uncertainties in scientific work that do not lead to immediate adjudication of some apparent discrepancies." *

---

* From pages 42-46 of the transcript of the Subcommittee hearings.

More recently, the Panel itself was forced to conduct its review, with the benefit of hindsight, but without clear guidelines as to when and whether notification of suspected misconduct should be given to the NIH. By documenting the experiences of the Brigham and Women's Hospital in these unfortunate circumstances, the Panel's investigation serves to focus the issues facing the NIH and scientific institutions and provide data for prospective guidelines.

C. Brigham and Women's personnel have taken the initiative in investigating Dr. Darsee's misconduct.

The Report acknowledges that Dr. Kloner and I investigated allegations of irregularities in Dr. Darsee's experimentation when they first came to our attention, in May 1981, immediately confronted Dr. Darsee with the evidence of misconduct we had discovered, obtained his admission to the single incident of fabrication of research data, promptly notified Deans Tosteson and Adelstein of the Harvard Medical School and Dr. Robert Petersdorf (then President of the Brigham and Women's Hospital) of the fabrication admitted by Dr. Darsee, and continued our investigation thereafter. Moreover, the Report confirms that, after the NHLBI established its own Panel in December 1981: "At all times, Drs. Tosteson, Braunwald, Kloner, and all other officials and employees at Harvard and the Brigham and Women's Hospital cooperated fully and promptly in providing all requested information and in facilitating the Panel's study in every way possible." (Page 32.)

As the Report describes, our efforts were not limited to cooperation with the Panel. Rather, upon discovering a single instance of data fabrication, Dr. Kloner and I took the initiative and "began an audit of all of Dr. Darsee's data and publications from the Cardiac Research Laboratory." (Page 11.) In October 1981, as a result of discoveries in the course of Dr. Kloner's and my continuing analysis of AMPIM data supplied by the NHLBI, after discussions

with co-investigators we advised the AMPIM coordinating committee that we "could not vouch for some of the data collected by Dr. Darsee," informed them about Dr. Darsee's prior admitted misconduct and "offered to repeat the AMPIM study." (Page 13.) We subsequently reported that it was "[d]iscrepancies between pre-May 22 and post-May 22 results in Dr. Darsee's blood flow and heart weight data" which led us to our conclusion that Dr. Darsee's misconduct extended beyond the single admitted fabrication and tainted his AMPIM data. (Page 13.) In the course of our continuing investigation, Dr. Kloner and I discovered and reported to the Panel evidence "that many of the dogs on which Dr. Darsee had reported regional myocardial blood flow were not included in the radioactive disposal log" (page 15; also page 23), suggesting that Dr. Darsee may not have injected microspheres into these dogs and, therefore, had not actually performed blood flow measurements upon them. Moreover, the Report concludes, "Drs. Braunwald and Kloner deserve full credit for initiating the measurement of residual radioactivity in cardiac tissues. This technique led to the important discovery that there was an absence of residual radioactivity in many of the hearts in which Dr. Darsee purported to have measured myocardial blood flow by injecting radioactive microspheres." (Page 23.)

Despite its belief that its own charge was limited to investigating Dr. Darsee's activities at the Brigham and Women's Cardiac Research Laboratory (page 3), the Panel notes that I extended my investigation to other settings in which I had reason to suspect misconduct by Dr. Darsee. Having discovered irregularities in a manuscript submitted for publication by Dr. Darsee with Dr. Leonard Holman of Harvard's Department of Radiology, I called these to the attention of Dr. Holman, who immediately withdrew both papers which Dr. Darsee had co-authored. (Page 13.) Identifying and purging the full extent of the misconduct which, when the researcher has worked in more than one laboratory

and institution, may well span institutional boundaries, would seem to be among the most important responses to an instance of dishonesty in research. Consequently, although the Report omits this point, I have, since November 1981, repeatedly urged appropriate personnel at Emory University School of Medicine to conduct an audit of Dr. Darsee's work during his five years at Emory and also urged the Panel to address the validity of Dr. Darsee's data at Emory.

D. Supervision and operation of the Cardiac Research Laboratory are adequate.

The Panel's Report concludes that: "The [Cardiac Research] Laboratory has had an outstanding record of productivity since its establishment by Dr. Braunwald in 1973 [sic: 1972]. It has provided research training and experience for many young investigators." (Page 27.) "In many respects supervision in the Cardiac Research Laboratory . . . has been adequate and comparable to that provided in similar laboratories at other institutions." (Page 28.) In addition, the Report notes that "as a consequence of Dr. Darsee's misconduct, this laboratory has undergone an unusually rigorous scrutiny," and "appropriate changes have been implemented by Drs. Braunwald and Kloner to improve the overall operation of the laboratory." (Page 33, also page 28.)

Although the Report is critical of some of our actions, it does not assert that the course it deems preferable should have been intuitively obvious to us at the time we made our decisions. Nor does the Report assert that there was, at the time, common agreement throughout the scientific community on how such an act of misconduct should be handled.

Dr. Darsee's admission of a single instance of misconduct thrust me and the other responsible officials into a gray area of the scientific enterprise where, as the late Dr. Handler testified before the Congressional committee on oversight, we have "no understandings among ourselves as to how any one such incident

shall be treated." Acting upon imperfect information, we honestly and diligently sought to balance fairness to Dr. Darsee with the need to ensure the integrity of the laboratory's work. The fact that the Panel believed that other actions would have been preferable to some of those we settled upon should not reflect unfavorably on the laboratory or the institutions; it simply illustrates that reasonable persons can arrive at different conclusions when forced to make such judgments ad hoc.

The Panel's review of these events underscores the needs for standards of general applicability. Judgments made in any particular investigation do not necessarily provide prospective guidance in future situations presenting different facts. In its Report, the Panel emphasized that the judgments it made were the product of "the altered perspective afforded by retrospective view." (Page 19, also pages 20, 32, 33.) Hindsight is, of course, an unfair vantage from which to penalize persons and institutions acting in good faith in the absence of clear standards, but it is a good perspective from which to fashion such standards.

II. Specific Comments

These comments are divided into seven areas:

A. The charge to the Panel (or the Panel's interpretation of its charge) inappropriately limited investigation to the Cardiac Research Laboratory.

I respectfully request that the NIH examine the charge to the Panel. Although the Panel interpreted its charge to include investigation of Dr. Darsee's activities and laboratory practices only during his association with the Cardiac Research Laboratory, there is reason to suspect that he engaged in data falsification at the two other laboratories in which he had previously

worked. By excluding consideration of the full range and circumstances of Dr. Darsee's possible misconduct, the investigation inappropriately and inaccurately focused the problem on our laboratory alone. Moreover, this limitation excluded comparative evidence that was essential to a realistic and meaningful evaluation of our laboratory's procedures and responses.

In responding to what I considered to be my scientific obligations, I personally analyzed Dr. Darsee's research data from each of the other two laboratories in which he had worked. I am convinced, and the findings of the Ad Hoc Advisory Committee to the Dean of the Harvard Medical School on Dishonesty in Scientific Research and of the NHLBI Panel have confirmed, that Dr. Darsee engaged in fabrication of research data in the Nuclear Medicine Research Laboratory at Brigham and Women's Hospital under the supervision of Dr. Leonard Holman, Professor of Radiology at Harvard. I reported my findings to Dr. Holman and he withdrew the papers in question. I have also examined several of Dr. Darsee's published papers from Emory University, where he worked for five years immediately prior to coming to Harvard. On the basis of my review, I have serious questions about the validity of some of the data in papers from Emory (Appendix A), some of which were sponsored by the NIH, and I have reported these questions to Dr. Darsee's superiors at Emory, and to the Panel and to NHLBI staff.

I do not question the propriety of investigating Dr. Darsee's conduct in the Cardiac Research Laboratory, but do question the fairness of singling out for the brunt of criticism the one laboratory of the three which (1) discovered his cheating in the first place; (2) accepted the responsibility to analyze his data from the other laboratories; and (3) reported to the others and the NHLBI the findings that his data reported in other laboratories were suspect. It should also be noted that Dr. Darsee's research at Emory and in the Nuclear

Medicine Laboratory involved patients, whereas his work at the Cardiac Research Laboratory was limited to dogs.

Most importantly, I believe that criticism of our practices and supervision cannot fairly be based on Dr. Darsee's ability to cheat in our laboratory unless the NIH has examined the practices and supervision in the other laboratories where he may also have cheated. Only if practices and supervision in the other laboratories were found to be inferior could Dr. Darsee's ability to cheat be related to the practices and supervision in the laboratories in which he worked. If practices and supervision in the other laboratories were judged to be normal and routine, as I firmly believe ours were, and as I expect that they were in the Nuclear Research Laboratory and at Emory, then the only conclusion justified by this evidence is that a clever researcher determined to cheat will get away with it, for a time, even under optimal conditions.

B. The Report's suggestion that our investigation of Dr. Darsee's work was insufficiently rigorous ignores the extraordinary difficulty, illustrated by the Panel's own experience in this matter, of investigating possibly fraudulent research.

The Panel established to investigate Dr. Darsee's misconduct consisted of a group of distinguished scientists which had available to it substantial resources and a dedicated and competent NHLBI staff. The Panel had access to all of Dr. Darsee's research materials, tissue specimens and notebooks, as well as all laboratory records, and had full opportunity to interview all relevant parties. In fact, Dr. Kloner brought the remaining AMPIM heart tissue blocks to Hershey, Pennsylvania on March 18, 1982 for examination by

members of the Panel and staff. At the Panel's request, the staff conducted in-depth statistical analyses of the AMPIM data and Dr. Darsee's published work. The statistical evidence developed as a result indicated broader misconduct by Dr. Darsee and raised serious questions about his data. Nevertheless, over the five months of its work the Panel and its staff never suggested measuring residual radioactivity in the remaining heart tissue, and therefore did not uncover the startling fact and most powerful evidence of misconduct that Dr. Darsee did not inject radioisotopes for the measurement of blood flow in the 35 AMPIM dogs he operated on before May 22, 1981 and in a number of dogs from his published work.

Instead, it was during our own investigation of the radioactive dog disposal log that, in April 1982, Dr. Kloner and I identified the possibility that Dr. Darsee had failed to inject radioactive microspheres into dogs on which he reported measuring radioactive myocardial blood flow. We reported this finding immediately to the Panel and, subsequently, recognized that by measuring residual radioactivity in the remaining AMPIM heart tissue blocks and by microscopic examination of these tissues we could determine whether Dr. Darsee had injected radioactive microspheres into the AMPIM dogs on which he claimed to have measured blood flow. It was the results of our measurements, which we reported immediately to the Panel, and which have since been confirmed by independent investigators, which constitute the strongest proof yet developed of the wide extent of Dr. Darsee's misconduct.

If our laboratory had <u>not</u> made these findings, and the matter were reviewed at some time in the future, the Panel's failure to examine remaining tissue for radioactivity and microscopic evidence of spheres--a procedure that seems obvious in retrospect--might have exposed the Panel's own investigation to criticism, in retrospect, as having been less than rigorous.

The point is not that the Panel lacked rigor, expertise or commitment in its investigation. Rather, this example demonstrates the inherent difficulty of investigating fraudulent scientific data and uncovering conclusive as opposed to circumstantial evidence of misconduct. This difficulty is compounded when the perpetrator refuses to admit misconduct and insists that his research is accurate. We are not trained, any more than the distinguished Panel members are, to investigate scientific fraud. Our failure initially to turn up evidence of Dr. Darsee's widespread misconduct, like the Panel's failure to inquire into the presence of radioactivity or of microspheres in the residual samples, does not mean that either audit was less than rigorous.

Furthermore, it should be noted that at the time of the Panel's visit in March, 1982, Dr. Kloner and I had by no means completed our investigation of the matter. As reviewed in detail on page 4 of Appendix F, we harbored serious suspicions (which we discussed with the Panel on March 3, 1982) that Dr. Darsee may have unblinded the AMPIM experiments and we were repeating many of his studies. In my written statement of March 3, 1982, prepared and delivered to the Panel prior to learning the Panel's concerns or statistical findings, I indicated that Dr. Kloner had repeated several key experiments. We believed that repetition of the experiments provided the best method of confirming the results of Dr. Darsee's experiments. Furthermore, I reported to the Panel that Dean Tosteson and I had agreed that he would appoint a committee of senior faculty to examine carefully these reviews and experiments. The Dean had, in fact, by the time of the Panel's first meeting with us, asked Drs. Barger and Benacerraf to serve in this capacity.

In retrospect, I agree that the better policy is to have uninvolved investigators perform such an audit. However, in view of the possible extreme difficulties inherent in investigating alleged research fraud, as demonstrated

by the Panel's experience in this case, I repectfully request that our efforts and ultimate discoveries be placed in a more positive light. As Dr. Fredrickson stated in testimony before the Congressional committee on oversight, the investigation and discrediting of a scientist for research fraud "moves very deliberately, because one must be sure. . . .[E]xclusion from science is the ultimate penalty, and one must proceed carefully, because there are ambiguities, uncertainties in scientific work that do not lead to immediate adjudication of some apparent discrepancies." References in the Panel's Report to the rigor of our audit, and speculation that uninvolved investigators in May 1981 might have discovered the depth of Dr. Darsee's misconduct sooner than we ultimately did, should be rejected by the NIH and not included in any published report.

C. Suggestions in the Report that the quality or quantity of supervision in the laboratory enhanced Dr. Darsee's ability to cheat or avoid detection are unjustified by the Panel's limited investigation.

The Panel interpreted its charge as limiting its investigation of supervision in the laboratory to the years 1979 to the present and only to Dr. Darsee's work in the Cardiac Research Laboratory at Harvard. The effect of this restriction was to consider this laboratory's procedures in a vacuum and ignore at least three important areas of inquiry: 1) the unblemished record of my laboratory prior to July 1979; 2) evidence of supervisory practices at comparable laboratories; and 3) Dr. Darsee's ability to fabricate data in an entirely separate laboratory at Harvard and possibly at a different University. As a result, the Panel had no comparative evidence concerning the relationship between supervision in the laboratory and Dr. Darsee's ability to cheat. I believe that it is inappropriate for the Panel to infer from Dr. Darsee's misconduct in the Cardiac Research Laboratory that the quality of supervision in the laboratory was somehow responsible for the ease with which he was able to produce such

fabricated data, especially when he may also have fabricated data in two other entirely separate laboratories in which he had worked.

The Report explicitly concludes that:

> "General supervisory practices [in the Cardiac Research Laboratory] have been based, appropriately, on an assumption of honesty and trust between supervisors and trainees. With the exception of Dr. Darsee, recent trainees have retained original records in customary and acceptable fashion. They have been allowed appropriately increasing independence as their laboratory experience has progressed. Because most incoming trainees have had superior records of previous accomplishment, they have perhaps been capable of progressing more rapidly than trainees in many other laboratories." (Page 28).

The supervisory practices employed in my laboratory have been designed not only to ensure the reliability of results but also to encourage initiative. Without wishing to be immodest, I do not believe that any other laboratory has produced more leaders in academic cardiology. In November 1981, I was honored to receive the Herrick Award from the American Heart Association. The citation includes the following statement: "His gift for training has favorably influenced many researchers and more than 40 former trainees are full professors, department heads or directors of cardiology divisions in our medical schools." (Appendix B)

I believe that the supervisory practices in my laboratory since 1960 have been similar to, and would compare favorably with, those of other good cardiovascular laboratories in clinical departments, were such evidence considered. In order for the Panel to get a broader base of information from which to make a judgment about this important matter, I specifically requested the Panel by letter of April 12, 1982, and again on June 17, 1982, to interview two former Fellows currently at the same institution as one Panel member. (One of these two former Fellows overlapped Dr. Darsee in the Cardiac Research Laboratory.) There is no indication in the Report that this was done.

Our laboratory procedures have, I believe, been measured against a stricter and more detailed standard than that applied to any other similar laboratory.

At the same time, the quality of supervision in our laboratory has not been compared with the quality of supervision in any other comparable laboratory. Therefore I believe that making absolute judgments and subjective criticism about the quality of our supervision is unjustified. I believe that it would be hazardous and ultimately perhaps even a disservice to science for NIH to make public subjective criticisms of individuals and institutions without a full and fair hearing, and a complete investigation of all relevant facts.

I have separated out each statement regarding supervision and laboratory procedures (Section VIII-H of the Report) to which I wish to respond. I bring these issues to the attention of the NIH in light of the statement in Dr. Raub's letter to me of September 10, 1982 that among the range of possible NIH actions was "formal NIH acceptance--or rejection--of particular conclusions of the Panel with regard to laboratory procedures, supervision, . . . and other issues of professional competence raised by the Panel's Report."

1. Report: "The review of data included in abstracts and papers to be submitted for publication . . . often relied on the ongoing examination of raw data and tracings conducted at an earlier time." (Page 27.)

The Report correctly notes that in some instances the results of all experiments were carefully analyzed on an ongoing basis, and in others the results of all experiments were reviewed prior to the preparation of an abstract or paper. However, I believe that the critical factor on which to focus is not the timing of the review, that is, whether it occurred shortly after the experiment, shortly before preparation of a paper, or both, but rather that a careful review of the data was always carried out. Indeed, the initial discovery of research misconduct which Dr. Darsee acknowledged in May, 1981 came about not as a consequence of the review of his research on an ongoing basis, but as the result of a request by Dr. Kloner to review all of his tracings prior to the preparation of an abstract, a point the Report did not comment upon. The

specific timing of this review would make no significant difference in the ability of the supervisors to uncover misconduct by a researcher determined to mislead them.

2. Report: "Log books were not carefully maintained in the laboratory . . ., the fellows' individual notebooks being regarded as the primary record." (Page 27.)

It is, of course, desirable to maintain the laboratory log book carefully. As the Panel noted, however, the Fellows' individual notebooks rather than the laboratory log book, were the primary data record in the laboratory. The Fellows' individual notebooks contained all data relating to the individual experiments, whereas the log book was simply one way for the laboratory in a single log to keep track of the identification number of each dog, the name of the Fellow operating on the dog, the date and the experiment to be performed. It should also be noted that the Panel acknowledges that the Fellows in the laboratory currently maintain the laboratory log book adequately. (Page 27.) Moreover, scrupulous insistence that the researchers regularly record experiments in the laboratory log, as well as their own notebooks, would not have prevented Dr. Darsee from fabricating data and recording them dishonestly twice, rather than once. Indeed, during our continuing investigation of Dr. Darsee's research, we reported to the Panel in April that we had discovered evidence that Dr. Darsee had falsified entries in the laboratory's log book.

3. Report: Dr. Braunwald's "multiple responsibilities made it difficult for him to supervise the overall operation of the Research Laboratory in 1979-81. The clinical, teaching and administrative duties which he has assumed do not allow sufficient time for direct participation in the research laboratory." (Page 28.)

The Panel did not examine my activities in detail. The Panel did not inquire into the manner in which I manage my responsibilities outside of those relating to the laboratory or delegate portions of those responsibilities to

capable colleagues. For example, the Panel did not inquire and, insofar as I know, was not aware that I undertook the responsibilities for consolidating Harvard's Department of Medicine at the Beth Israel and Brigham and Women's Hospital on the condition that a clinical head for each medical service would be named, to leave me time to continue scholarly activities. Moreover, this responsibility did not commence until July 1980, a full year after Dr. Darsee came to the laboratory and eight months after his dogs started failing to show up on the radioactive dog disposal log, suggesting that he began cheating long before I assumed my duties at the Beth Israel. Yet the Panel apparently considered these duties in implying that I might have had had too many responsibilities to manage the laboratory so as to prevent Dr. Darsee's misconduct or to catch it earlier.

The Report also does not note that the past sixteen months have been entirely unrepresentative insofar as my activities in the laboratory are concerned, since the matter concerning Dr. Darsee's misconduct has taken up an enormous amount of my time. These factors have not permitted me to devote nearly the same effort to the laboratory that I had during the preceding 13 years (1968-1981) when I had also been a Department Chairman. Certainly, the effect of the Darsee investigation upon my participation in the laboratory's activities during this period was reflected in the interviews of fellows in the laboratory, on which the Panel based many of its conclusions in this area. Nevertheless, even during this unusual period, I have continued to meet twice weekly with Dr. Kloner to review all pertinent laboratory data, and I have maintained major involvement in the development of research ideas and the writing of abstracts and papers. This may not have been apparent to the Fellows with whom the Panel spoke.

Some of the Report's criticisms stem from an emphasis on the question of my day-to-day physical participation in the laboratory. One of the key roles of a Principal Investigator is to provide scientific leadership, ideas and, when possible, even inspiration. Since 1960 I have maintained my direct participation in the laboratory by trying to offer these while at the same time providing an extremely capable colleague, such as Dr. Kloner, for day-to-day, on site, full-time participation in the actual experiments. I respectfully submit that this is the norm for Principal Investigators who are senior members of a faculty. Focusing on the manner in which I discharged my responsibilities as Principal Investigator without examining in comparable detail the manner in which other senior faculty discharge theirs is unfair and unrealistic. In any event, unless Dr. Kloner or I had stood over Dr. Darsee for the entire duration of his experiments, it is unlikely that given this particular situation, we would have uncovered his misconduct sooner than we ultimately did.

Supervision is clearly necessary, but no amount of direct participation by senior scientists can foreclose the possibility of intentional misconduct by a researcher. On the other hand, excessive direct supervision could chill researchers' personal initiative and independence of thought, essential ingredients to the success of any laboratory. Indeed, the actions of researchers and technicians in the laboratory who developed and reported their suspicions about Dr. Darsee's research practices are attributable, in part, to an environment that encourages independent initiative and dedication to science. Moreover, my laboratory's record of training successful and productive scientists demonstrates the benefits of such an environment.

I do not disagree that my "multiple responsibilities made it difficult . . . to supervise the overall operation of the Research Laboratory in 1979-1981." However, I strongly disagree with the possible implication, although the Report

does not so state, that because of the difficulty, the supervision was not carried out. That it was difficult does not mean it was not done. I believe that it was inappropriate for the Report to raise this subjective implication concerning my supervision of the laboratory without unambiguous evidence and without an opportunity for the matter to be openly presented, with the questions posed and responses evaluated. Moreover, to suggest a connection between my varied responsibilities and the ability of Dr. Darsee to cheat on his research in the Cardiac Research Laboratory without an examination of the methods and responsibilities of senior scientists generally or in the other two laboratories in which Dr. Darsee worked is unreasonable.

4. Report: Dr. Kloner "has made a genuine effort to discharge his duties in a careful manner but they appear to be too numerous and varied. He has not had an opportunity to assume these responsibilities in a gradual, phased manner." (Page 28.)

I believe that this contention is unjustified. Prior to assuming his responsibilities in 1979, Dr. Kloner received training in cardiovascular research in three separate laboratories over seven years and earned both M.D. and Ph.D. degrees. He was awarded an Established Investigatorship of the American Heart Association. Neither Dr. Kloner nor the laboratory can be faulted on this account for Dr. Darsee's breach of trust.

5. Report: "[T]he laboratory cannot be said to have performed optimally in ensuring appropriate handling, retention, and reporting of primary data after their collection by all fellows." (Page 28.)

The Report specifically concludes that other Fellows in the laboratory handled and retained their primary data in an acceptable and appropriate manner. (Pages 27, 28.) This criticism, therefore, relies solely on the laboratory's experience with Dr. Darsee. There is no question that Dr. Darsee misled Dr. Kloner and me consistently and that he provided a great deal of false information to us. It does not follow from this however, that "optimal" laboratory

procedures can prevent a determined and clever individual from temporarily concealing misconduct.

Because Dr. Darsee came to Harvard with such high recommendations, and was completing his seventh year of postdoctoral training at the time his misconduct was first discovered, it did not appear to Dr. Kloner and me to be necessary (or indeed desirable) to scrutinize his every activity, or to recheck all of his primary data. Dr. Darsee was observed by Dr. Kloner and others to be carrying out experiments in the laboratory, analyzing records and calculating results. He was a prodigious worker who stayed until late at night in the laboratory several times a week and on weekends. He showed and discussed the results of individual experiments with Dr. Kloner almost on a daily basis. Then, at a series of research meetings with Dr. Kloner and me, Dr. Darsee presented the tracings and graphs of results of several experiments. These were reviewed in the meetings and plans for further experimentation were developed by Drs. Kloner, Darsee and myself. Subequently Dr.Darsee was observed to be carrying out experiments by Dr. Kloner and returned to Dr. Kloner and me with what appeared to be accurate results.

The laboratory was operated on the general assumption that if a Research Fellow were observed by Dr. Kloner and others to be diligently carrying out experiments, analyzing data daily, and if the results presented at data meetings generally matched those observed by Dr. Kloner during the course of an experiment or in his review of the results at the end of the day, there was no reason to question the veracity of the researcher or his data. For example, if: 1) At a data review meeting Dr. Darsee had shown Dr. Kloner and me hemodynamic tracings which Dr. Darsee asserted he had obtained from four animals; if 2) it was agreed by Drs. Kloner, Darsee and myself that another four experiments were necessary to provide an adequate test of the hypothesis under

study; if 3) Dr. Darsee were then observed by Dr. Kloner to continue to work regularly on the experiment in question; and if 4) he brought the additional tracings and results to the next meeting several weeks later (often with suggestions for figures, tables and diagrams, as well as a review of the literature) neither Dr. Kloner nor I doubted the veracity of the material.

Dr. Darsee rapidly became more adept in experimentation and assumed responsibility commensurate with his increasing skills. That is not to say that he was unsupervised, but, for example, once it was clear to Dr. Kloner and me that Dr. Darsee could accurately measure ventricular end-diastolic pressure or the velocity of myocardial shortening, an increasing proportion of the time we spent with him was devoted to discussing experimental design.

6. Report: "A hurried pace and emphasis on productivity" existed in the laboratory. (Page 28.)

The pace in the laboratory from 1979 to 1981 was no different than the pace in my laboratory over the preceding 20 years, when there were no instances of research misconduct. The vast bulk of the research in the laboratory was supported by NIH and the pace of the laboratory's efforts, which I believe was not at all unusual, was dictated by the responsibility that we felt to live up to commitments made in the grants and contracts which supported our work. This assertion in the Report is entirely subjective and unsupported by either careful examination of projects being conducted in the laboratory or, more importantly, by comparison with the paces of comparable laboratories around the country.

7. Report: Research Fellows had limited interaction with senior scientists." (Page 28.)

As already stated, the Research Fellows met almost daily with Dr. Kloner and at intervals with me. I deliberately provided much of my input through my biweekly meetings with Dr. Kloner, and not in face-to-face meetings with the individual Fellows. In working for many years with scientists in positions of intermediate seniority such as Dr. Kloner, I have learned to make every effort not to appear to undercut their authority and responsibility for dealing directly with the Fellows.

8. Report: "The Panel was told of one published paper in which some controls from a previous published paper by the same investigator were used, with Dr. Kloner's approval, without the text so specifying," (Page 28.)

The practice of not specifying in the Methods Section that a control dog had been used in more than one study, while certainly not ideal, is common. The use of control dogs in more than one study was not done to mislead, nor did it cause any biased results (except in Dr. Darsee's case). This issue is discussed more fully in Appendix F, at pages 10-I2.

9. Report: "The supervisory practices in the laboratory, while in no way responsible for Dr. Darsee's misconduct, may have contributed inadvertently to the ease with which he was able to produce fabricated data and to the subsequent difficulty in documenting the extent of the problem." (Page 28.)

For the reasons set forth in pages to 15-17 above, I believe that this summary statement is unjustified.

Several of the criticisms in the Report concerning aspects of the supervision in the laboratory are subjective, casual references unsupported by a detailed examination of my role in the laboratory or supervision in comparable laboratories. None of the Report's preferred procedures would have prevented Dr. Darsee from fabricating research data and escaping detection for a time. Only a rigid system requiring a researcher's superiors to supervise personally the conduct and recording of every aspect of each experiment would effectively

prevent deliberate fraud, and nothing in the Report recommends such an impractical procedure.

I believe that the Report's statements concerning my varied professional responsibilities and the pace in the laboratory should be rejected by the NIH and not included in a published NIH Report. Before such speculations and impressions are publicized in a government report, the NIH should establish a standard in which the burden of producing hard evidence to support generalized or subjective criticisms is borne by the party making the accusations. The Report, as it now stands, fails to provide evidence to support these several subjective assertions. In the absence of a full hearing with the allegations specified and an opportunity provided for a response, such assertions simply should not be made.

A full and fair opportunity to test the evidence and conclusions on this point was never provided by the Panel. The first time that I was aware that the Panel intended to comment on my own role in the laboratory and my other responsibilities at the University and the Hospital was when I was afforded an opportunity to review the Panel's Draft Report. The Draft Report contained statements similar to those contained in the Final Report regarding my role in the laboratory and my varied responsibilities. I responded in writing to those statements and at the same time requested the opportunity to appear again before the Panel to discuss the matter of supervision in more detail. I regret very much that the Panel refused to provide me this opportunity.

In December 1981, a Subcommittee of the Ad Hoc Advisory Committee to the Dean of the Harvard Medical School on Dishonesty in Scientific Research conducted a site visit of the laboratory and, in the words of their report, "reviewed the systems for data collection and preservation used by Dr. Darsee and other members of the laboratory." The Committee concluded that:

"(t)he Subcommittee visiting the laboratory and the Committee as a whole were convinced that the laboratory directed by Dr. Kloner was well administered with appropriate regard to data collection and review. The visiting Subcommittee noted that data books dating back to 1977 revealed an 'impressive set of well maintained, annotated and stored raw data that was in good enough condition to allow reanalysis even at this time and seemed complete with protocol, mathematical analysis and finished conclusions.' The Subcommittee concluded that: 'the present problem does not appear to be at all referable to the existing Cardiac Research Laboratory standards, policies or procedures, nor to overt pressure provided by its director or Dr. Braunwald. Dr. Robert Kloner and the Cardiac Research Laboratory have maintained an extremely effective system for data collection, analysis and storage.'"

The NHLBI Panel in its Report reviewed the principal conclusions of the Ad Hoc Committee and, although it explicitly disagreed with another conclusion of the Committee, it did not do so with the Committee's conclusions, stated above, regarding the standards, policies and procedures in the laboratory. Rather, in the face of the Ad Hoc Committee's conclusions, without a thorough examination of the evidence (which would include an analysis of supervisory practices in comparable laboratories in clinical departments, as well as the other laboratories in which Dr. Darsee worked and produced suspect data), and without a full and fair hearing, the Panel offered subjective conclusions which, if accepted by the NIH, could affect past and present, honest and dedicated scientists in the laboratory. For the several reasons stated above, I respectfully request that all findings in the Report relating to my supervision in the laboratory be rejected and not included in a published report.

All of the above does not imply, however, that there were not changes and improvements which could be made in the procedures and supervision in the Cardiac Research Laboratory, as in any laboratory. Changes and improvements have been made and these have been reviewed by the Panel. As stated above, we appreciate the Panel's confirmation that "appropriate changes have been implemented by Drs. Braunwald and Kloner to improve the overall operation of the laboratory." (Page 33.)

D. The Report does not state the complex considerations we faced in May 1981 which led us to make the decisions we did.

The Report concludes that "[w]hen the acknowledged episode of data falsification by Dr. Darsee was uncovered in May 1981, Dr. Braunwald acted promptly to terminate Dr. Darsee's Harvard appointment and NIH Fellowship. He appropriately informed administrative officials at Harvard and the Brigham and Women's Hospital of the circumstances of the incident." (Page 32.) However, the Report questions other aspects of our response at that juncture, including our decisions not to notify the NHLBI immediately of the May discovery, to permit Dr. Darsee to continue to perform AMPIM experiments under closer supervision, and to undertake an internal rather than external audit of Dr. Darsee's work.

Of course, had I known then what I now know about the extent of Dr. Darsee's misconduct, I would have reported it immediately to the NIH and insisted upon an external investigation. However, in criticizing our decisions on these matters the Report does not give appropriate weight to the complex and competing considerations we faced in May 1981. Most significantly, as the Report acknowledges, there were no guidelines from the NIH (or, indeed, from any other body) to follow in responding to the incident. The report does not fully credit the severe sanctions we imposed against Dr. Darsee and does not deal with the legal, ethical and scientific factors which we considered in May 1981 as we responded to the initial discovery of a single incident of admitted misconduct. Because these considerations are not described in the Report, the failure to notify the NHLBI and the decision to permit Dr. Darsee to continue to perform AMPIM experiments may seem inexplicable.

It was my concern that Dr. Darsee be treated fairly, in the absence of relevant guidelines, that led to the decision to take no action in May to publicize

the facts surrounding his dismissal from the Hospital and Harvard. Dr. Fredrickson, in his comments to the Congressional committee on oversight, placed these concerns in clear relief when he stated "we can destroy them for a whole career. We can cast them out of science and one has to be extraordinarily careful in the exercise of judgment. . . ."

In summary, the situation we faced in May was the following: Dr. Darsee had come to Harvard Medical School and Brigham and Women's Hospital as a Research Fellow in Medicine in July 1979. He had received the highest possible recommendations from senior faculty at Emory University where he had served for five years in a variety of capacities, including Chief Resident in Medicine and Clinical Fellow in Cardiology. His performance during the first 22 months of his Fellowship at Harvard Medical School/Brigham and Women's Hospital was perceived as superlative in every respect and he was offered an Assistant Professorship at Harvard Medical School. In May 1981, Dr. Kloner caught Dr. Darsee fabricating tracings in connection with the preparation of an abstract for submission to the AHA. Dr. Darsee insisted that he had performed the experiments in question, but stated that he was unable to locate the original tracings and, when requested by Dr. Kloner to review the original tracings with him before the abstract was submitted, he had fabricaterd tracings to replace those that he had originally taken and subsequently misplaced.

Dr. Kloner promptly reported the incident to me, and I immediately confronted Dr. Darsee with the evidence concerning his misconduct. When Dr. Darsee admitted recreating the tracings in question, I took the following actions: 1) I immediately withdrew my offer of an Assistant Professorship and terminated Dr. Darsee's fellowships at Harvard Medical School/Brigham and Women's Hospital and his Postdoctoral Fellowship supported by the NIH. 2) I immediately notified Deans Tosteson and Adelstein of the Harvard

Medical School and Dr. Robert Petersdorf, then President of Brigham and Women's Hospital, orally and in writing. 3) Dr. Kloner and I immediately began a detailed analysis of Dr. Darsee's research work in the laboratory. Our review included studying his laboratory notebooks and interviewing each person who had worked in the laboratory with Dr. Darsee about any suspicions he had of possible misconduct. This initial investigation did not turn up anything to support a finding of additional misconduct. 4) I permitted Dr. Darsee to continue working in the laboratory in an informal capacity under increased supervision to assist in the audit and for two and a half months to complete some experiments on which he had been working, and so that he could be observed carefully as the situation became clarified to determine whether additional actions were warranted. 5) In the Fall of 1981, Dr. Kloner and I notified two prospective employers of Dr. Darsee's single act of scientific dishonesty.

What we knew then, in May 1981, was that an apparently brilliant researcher had committed and acknowledged a single incident of misconduct. Our initial investigation uncovered no evidence of further misconduct. The researcher insisted that this act was a single incident and, indeed, that the original experiments in question had been performed and the fabricated tracings were only a recreation of the original tracings which he had misplaced. Even as we considered what actions to take, I was receiving letters from his former supervisors recommending Dr. Darsee's appointment as an Assistant Professor at Harvard Medical School and praising him highly. (Appendix C)

In light of the absence of guidelines or regulations from the NIH or any other source, I determined that there were essentially three kinds of responses available at the time. At one end of the spectrum, Dr. Darsee could have been issued a severe reprimand, placed under observation and continued at

the University, the Hospital and the Laboratory on probation. At the other end of the spectrum, we could essentially have terminated his academic career. He could have been dismissed from the University and from the Hospital and his single, admitted incident of misconduct could have been made public by, among other such steps, reporting the conduct to the NIH. I was not aware (and am still not aware) of channels through which the matter could have been reported to the NIH under an assurance of confidentiality. The middle ground consisted of notifying appropriate University and Hospital authorities immediately, as well as potential employers who made inquiry, withdrawing the offer of a faculty appointment, terminating his fellowships at Harvard, the Hospital and NIH, and beginning an internal audit of his work, but allowing him to complete experiments in the laboratory for a brief period, but under closer supervision.
This middle course seemed the most reasonable because it appeared at the time that only a single acknowledged act of recreating original tracings was involved. This decision was a difficult one, but in the absence of guidelines suggesting a different course, it appeared to balance the competing and complex legal, ethical and scientific interests in the fairest way by providing Dr. Darsee with due process while ensuring the scientific integrity of the research and data emanating from the Cardiac Research Laboratory.

When it became evident in late October that the scope of Dr. Darsee's misconduct was broader than initially known, I immediately took a series of more severe steps. Upon Dr. Kloner's discovery on October 29, 1981 that Dr. Darsee's AMPIM results for dogs studied before May 22, 1981 differed significantly from his results for dogs studied after that date, we immediately notified the NHLBI and the AMPIM investigators of Dr. Darsee's May misconduct and stated that his AMPIM results could not be vouched for. The oral notification was followed immediately by a written statement that we could not stand

behind Dr. Darsee's values produced for the AMPIM study and that the data should be withdrawn from the study. In addition to notifying of the NHLBI, I took numerous other steps at this time, including: 1) On November 6 all abstracts on which Dr. Darsee was senior author and which had been accepted for presentation at the American Heart Association meeting were withdrawn. 2) I asked Dean Tosteson to appoint a committee of distinguished faculty members, some from outside Harvard, to review the situation and devise guidelines and recommendations. 3) Upon learning that Dr. Darsee had continued to work (under false pretenses) on several experiments in a separate laboratory at Harvard under the direction of Dr. Leonard Holman, Professor of Radiology, I reviewed two manuscripts based upon those experiments. Finding the data suspicious in one of them, I urged Dr. Holman to withdraw the articles, which he did. 4) I contacted officials at the Emory University Medical School and informed them of the situation, recommending that they conduct an investigation of Dr. Darsee's work there. Thus, in October, when I learned that the extent of Dr. Darsee's misconduct went beyond what was originally known, I followed the third and most severe in the above-described spectrum of responses.

I believe that the statement of the Ad Hoc Advisory Committee to the Dean of the Harvard Medical School concerning the course of action we chose over the period from May to October provides an appropriate assessment of our decisions. The Committee stated:

> "When considered in two phases, the responses of the institution seem reasonable in relation to the information available at the respective times. The first phase response was prompt, but limited, and seemed appropriate for what then appeared to be a single episode of aberrant behavior. The realization that there might be more than one episode of dishonesty appropriately precipitated the larger and more comprehensive second phase response in November and December.
>
> "The committee questioned the reason for the delay between the initial event and response in May and June and the full scale investigation in November. In retrospect, it is clear that Drs. Braunwald and Kloner felt

in May and June that they were probably dealing with a single bizarre act by a young man who had performed exceptionally well previously. In the light of this consideration, the plan selected in June seems to be reasonable in that it provided an opportunity to assess the extent of the damage and also to provide for a period of observation under supervision."

In hindsight, the decision to permit Dr. Darsee to remain in the laboratory and perform AMPIM experiments for two and a half months under closer supervision may appear to some to have been a mistake. However, the failure of the Report to present the explanation for that decision or, indeed, the benefits that accrued therefrom to the investigation of Dr. Darsee's misconduct, presents an unbalanced view of our reasoning and motivation.

As stated above, all AMPIM data gathered in all the laboratories had been submitted to the NHLBI in April 1981, where it had been reviewed, and we had not been informed that there were any problems with the Darsee data as a result of that April 1981 data review. Assuming that Dr. Darsee had not engaged in any misconduct concerning the previous AMPIM experiments, replacement of Dr. Darsee could jeopardize the study's results. Assuming that Dr. Darsee had engaged in misconduct in the AMPIM studies performed up to that date (a possibility which I then considered to be extremely unlikely) our component of the study was already invalid and beyond salvage. Ultimately, the results of this decision enabled Dr. Kloner and me to compare the results of his experiments pre-May 22 (dogs #1-35) and post-May 22, and to provide the basis for nearly all subsequent conclusions by us and by the Panel as to the scope of Dr. Darsee's misconduct. The Ad Hoc Committee Report, while recognizing, as do I, that as it turned out there were heavy costs to the decision, makes this same point:

"As the Committee looks back at the events of the summer and fall of 1981, it appears clear that it was helpful to allow Dr. Darsee to continue working in the laboratory because the comparison of the pre-May 22 and post-May 22 data in the 'Models Study' proved to be of great value in the final analysis. On the other hand, this course of action was not without cost. Dr. Darsee's presence in the laboratory environment after the events of May proved to be damaging to the morale and productivity of the laboratory."

The fact is that allowing Dr. Darsee to continue on the study and, in effect, to serve as his own control did not interfere with the integrity of the study. As the Panel found, on the contrary it allowed us (in October) to develop powerful evidence that the study had already lost its integrity because of Dr. Darsee's misconduct prior to May 22. In all fairness, any report published by NIH, in discussing this issue, should acknowledge this fact.

E. The Panel's Report does not, but most certainly should, state that it was Dr. Kloner's finding that led us to conclude in October 1981 that Dr. Darsee's AMPIM data were not reliable.

The Report creates the misimpression that the NHLBI discovered and disclosed to Dr. Kloner and me in October 1981 the evidence of irregularities in Dr. Darsee's AMPIM data. It should be noted, however, that the results of the AMPIM data were blinded until October 1981, and that the most powerful clue to uncovering Dr. Darsee's misconduct in the AMPIM emerged from our comparison of the results he reported on experiments conducted before May 22, 1981 and the results he reported after that date. We did not receive all of the results of the study from NHLBI until October 1981 and we could not compare Dr. Darsee's results with those from other laboratories until then.

Moreover, I have reviewed the events of October 1981, and Dr. Kloner's letter to me of February 1, 1982 (a copy of which I provided to the Panel) setting out the sequence of these events. I read Dr. Kloner's letter to Dr. Reimer shortly after it was written and Dr. Reimer concurred with the statements in the letter as to his understanding of the sequence of events in October. I believe that it is beyond dispute that it was Dr. Kloner who discovered the difference between Dr. Darsee's pre- and post-May 22 results and that it was on the basis of that finding, following previous mutual discussions with coinvestigators, that Dr. Kloner and I reached the conclusion, and immediately advised the NHLBI, that the integrity of Dr. Darsee's data could not be vouched for.

We agree that, after the results were unblinded, collated and made available in October, NHLBI staff and other AMPIM investigators gradually became concerned about inconsistencies between the data from the Brigham and from the other centers and that there were mutual discussions among the investigators about these matters. But there were also discrepancies initially noted between results from two of the other laboratories. Specifically, there was a significant reduction in infarct size/risk region with the drug Verapamil at the VA Laboratory, but this effect was not present in the Johns Hopkins results. This finding puzzled us and occupied a good deal of our attention in October. It was during the second or third week in October that Dr. Kloner spoke with Dr. Reimer and asked him to forward copies of graphs which he had drawn computing the ratio of infarct size to risk region versus subepicardial flow, hoping that they might help us resolve inconsistencies in the data. Dr. Reimer sent Dr. Kloner a handwritten version of the graphs of necrosis over risk region versus flow, and it was while studying this graph that Dr. Kloner discovered the lack of comparability between pre- and post-May 22 data from Dr. Darsee. Dr. Kloner immediately informed me of this startling discovery by long distance telephone, and we decided together that he would inform the NHLBI immediately that we could not stand behind the data. It was on the basis of that discovery that we were able to conclude that Dr. Darsee's AMPIM data were not reliable.

The Panel's Report, unfortunately, could be misinterpreted to imply that the NHLBI staff was primarily responsible for discovering Dr. Darsee's wrongdoing on the AMPIM data. In fact, we acted responsibly and in good faith to review Dr. Darsee's results and compare the data from the various centers as soon as they became available to us in October. Within hours of our recognizing the disparity between the pre- and post-May 22 results on October 29, 1981, we informed the NHLBI that we could not stand behind the data. We then offered

to repeat the experiments. The Report is misleading if it does not make this point clearly.

F. It is our hope that the laboratory will be authorized to repeat its portion of the AMPIM study unless, because of the lack of rigorous blinding at three of the four institutions, the AMPIM results are not to be used.

As I informed the NHLBI Panel when I first met with it on March 3, 1982, I hope very deeply that we will be authorized to redo the experiments. I am anxious that the results of this important project be compiled and published without any further delay and I am confident that our laboratory is entirely capable of performing the AMPIM experiments with the required professional competence. The Panel itself concluded that "[a]part from the professional misconduct of Dr. John Darsee, the integrity and scientific capabilities of the professional and technical staff of the cardiovascular laboratory at the Brigham and Women's dospital are of a high order." (Page 32.) Regarding areas of procedure and supervision questioned by the Panel, the the Report states that "[s]ince these unfortunate events, appropriate changes have been implemented by Drs. Braunwald and Kloner to improve the overall operation of the laboratory." However, if the AMPIM results are not to be used because of the lack of rigorous blinding at three of the four institutions, as described in the Report (page 30), then clearly repetition by us of our portion is pointless and wasteful of valuable resources.

G. Consistent with their obligations to the scientific community and as a precautionary measure, Drs. Braunwald and Kloner have requested the editors of Journals which have published articles contributed by Dr. Darsee from Harvard to publish retractions of articles or sections of articles which are suspect.

Attached to this memorandum (Appendix D) are copies of the letters submitted by me and/or Dr. Kloner to scientific journals in which papers were published from the Cardiac Research Laboratory for which research data was provided by Dr.

Darsee. These letters were sent to the journals as soon as the residual radioactivity and microsphere data were discovered by us and confirmed by independent investigators at Harvard. We submitted the letters in response to our obligations to science and to the scientific community and after consulting with our colleagues at the Harvard Medical School and the Brigham and Women's Hospital and with members of the NHLBI Panel. All efforts were made in the letters to provide, consistent with the need for accuracy and fairness to all involved parties, information as complete as possible to inform the scientific community of our opinion of the data and the validity of the conclusions in the papers. Each journal involved was requested to publish our letter at the earliest possible date, and we have been informed that publication of the letters will begin in October.

APPENDIX A

I have reviewed the following papers published by Dr. Darsee when he was a Fellow in Cardiology at Emory University and regrettably have encountered problems which are similar in many respects to those in research which he carried out in the Cardiac Research Laboratory at the Harvard Medical School/ Brigham and Women's Hospital. In the case of the first two papers, indicated below, the concerns involve a much (unbelievably) closer correspondence between measurements of cardiac output between two relatively inaccurate methods than would be anticipated from duplicate measurements using either method alone.

In the case of the third paper, the ability to achieve perfect 24-hour urine collections in every member of a family of 46, including an infant, exceeds by far that which would be anticipated. In addition, examination of the creatine clearances raises two serious concerns: 1) Patients with heart failure demonstrate clearances almost identical to those of patients without heart failure, an observation which is at variance with established findings; 2) there is a greater variability of creatine clearances within each group of patients than between the six groups of patients.

1. Darsee, J.R., Mikolich, J. R., Walter, P. F., Schlant, R. C.: Doppler cardiac output 1. Comparison of transcutaneous and juxta-aortic Doppler velocity signals with catheter and cuff electromagnetic flowmeter measurements in closed and open chest dogs. Am. J. Cardiol. 46: 607-612, 1980.

2. Darsee, J. R., Walter, P. F., Nutter, D. O.: Doppler cardiac output II. Noninvasive measurement of transcutaneous Doppler aortic blood velocity integration and M-mode echocardiography. Am. J. Cardiol. 46: 613-618, 1980.

3. Darsee, J. R., Heymsfield, S. B.: Decreased myocardial taurine levels and hypertaurinuria in a kindred with mitral valve prolapse and congestive cardiomyopathy. New Engl. J. Med. 304: 129-135, 1981.

I notified officials at Emory University of Dr. Darsee's misconduct in November 1981 and directed attention to the questions about these papers in various conversations with the Chairman of the committee at Emory University investigating the matter, and with the Chairman and the Executive Secretary of the NHLBI Panel in the Spring of 1982.

APPENDIX B

AMERICAN HEART ASSOCIATION

JAMES B. HERRICK AWARD

Presented to

EUGENE BRAUNWALD, M.D.

by the

COUNCIL ON CLINICAL CARDIOLOGY

The 1981 James B. Herrick Award is presented to Eugene Braunwald, M.D. of Harvard Medical School for 25 years of extraordinary achievements which have advanced the practice and the science of clinical cardiology.

His contributions are many but his work on limiting the ultimate size of myocardial infarction influenced greatly the clinical practice of cardiology. In 1967, Dr. Braunwald proposed that the size of infarction is not determined solely by the site of occlusion, but can be influenced profoundly by the balance between myocardial oxygen demand and supply in the affected area. He reasoned that this was a dynamic process and not an unalterable sequence beginning with myocardial ischemia and ending with necrosis.

In a series of brilliant animal experiments, now amply confirmed by other investigators, he demonstrated that infarct size following coronary occlusion could be reduced by a variety of interventions, including beta-adrenergic blocking drugs, coronary vasodilators, and anti-inflammatory com-compounds, administered singly and in combination.

These observations, translated by Dr. Braunwald and his colleagues into an improvement in the clinical management of acute myocardial infarction, arose from basic laboratory research into the determinants of myocardial oxygen consumption by Braunwald in the 1950s and early '60s. They represent an example of clinical investigation at its best. A clear thread is evident

from his most fundamental research on myocardial energetics through animal experiments assessing the disease process, to the conspicuous improvement in treatment of a major disease entity in the industrialized world.

Dr. Braunwald received the M.D. degree from New York University, trained at Mt. Sinai Hospital and then became a Fellow in the laboratory of Andre Cournand at Bellevue Hospital. He joined the National Heart Institute as clinical associate in 1955 and later was Clinical Director. In 1968, Dr. Braunwald became the first Chairman of the Department of Medicine at the new University of California Medical School in San Diego. In 1972, he became Harvard's Hersey Professor of the Theory and Practice of Physic, one of the oldest and most prestigious chairs of medicine.

His earliest research work dealt with hemodynamics of the normal and diseased heart. In studies with Sarnoff, Dr. Braunwald characterized the hemodynamic determinants of myocardial oxygen consumption and coronary blood flow, identifying the tension time index as a major determinant of myocardial oxygen consumption. He and his colleagues clarified the importance of Starling's Law as a major determinant of ventricular performance in man; and, he contributed to the initial description of, and then named, idiopathic hypertrophic subaortic stenosis.

Dr. Braunwald and his colleagues at the NIH studied cardiac dimensions under a variety of conditions in intact unanesthetized patients in whom radiopaque markers had been sewn to the surface of the heart at the time of corrective cardiac operations. Also, they developed techniques for characterizing myocardial force-velocity relations in intact unanesthetized man, and identified velocity of cardiac contraction as a major determinant of myocardial oxygen consumption. In addition, they performed some of the earliest studies on

beta-adrenergic receptor blocking drugs and identified reduction of cardiac tyrosine hydroxylase activity as the immediate cause of norepinephrine depletion in heart failure.

Outside the laboratory Dr. Braunwald's accomplishments have been exceptional. He has written and edited many books and monographs on heart disease and cardiovascular physiology. His landmark textbook, Heart Disease, was published in 1980. His gift for training has favorably influenced many researchers and more than 40 former trainees are full professors, department heads or directors of cardiology divisions in our medical schools.

He is the only cardiologist holding membership in the National Academy of Sciences, is a member of the Institute of Medicine of the Academy, and a Fellow of the American Academy of Arts and Sciences. He has served as President of the American Society for Clinical Investigation, the American Federation for Clinical Research, and the Association of Professors of Medicine. His awards include the Abel Award for Research in Pharmacology and the American Heart Association Research Achievement Award.

In this voluntary arena, he has performed unstintingly: 4 years as a national vice president of AHA, 5 years as chairman of the Publications Committee, 2 years as chairman of this Council, 4 years on the Board of Directors and 15 years on editorial boards of AHA scientific journals.

Recognizing his meritorious contributions, the Council is pleased and proud to present the Herrick Award to Dr. Eugene Braunwald.

November 17, 1981 /s/ J. Schoenberger
President

/s/ Robert O'Rourke,
Council Chairman

APPENDIX C

EMORY UNIVERSITY SCHOOL OF MEDICINE
EMORY UNIVERSITY HOSPITAL
1364 Clifton Road, N.E. Atlanta, Georgia 30322

CLINICAL RESEARCH FACILITY (404) 329-7258

May 21, 1981

Eugene Braunwald, M.D.
Beth Israel Hospital
Harvard Medical School
330 Brookline Avenue
Boston, Massachusetts 02215

Dear Dr. Braunwald;

I am writing to support the promotion of Dr. John Darsee to Assistant Professor of Medicine. Most of John's virtuoso performance is apparent from his extensive bibliography, an accomplishment that is truly remarkable in that during much of this time John was a housestaff officer.

My first contact with John was in 1977 during his cardiology fellowship. Dr. Nutter and I had been frustrated by our inability to find a cardiology fellow to complete a study on skeletal muscle changes in patients with cardiomyopathy. When John rotated through our service, he was asked to help out in this program. Although at that point he had not yet written any major publications, much to our amazement the project was completed, written up, and subsequently published in Circulation with apparent ease.

The story from that point on is quite remarkable. Despite an enormous committment to teaching and patient care, John continued to collaborate and produce first class scientific publications throughout his remaining 3 years at Emory. He was a stimulus to all around him, and his knowledge and teaching sessions were sought after by all of those around him. He is warm and friendly, and is a true humanitarian.

I have no doubt that John is one of the leading internist-cardiologists of his age group in the country.

Sincerely,

Steven B. Heymsfield, M.D.
Director, Division of Nutrition
Associate Professor in Medicine
Asst. Director, Clinical Research Facility

RECEIVED
MAY 26 1981
DEPT. OF MEDICINE, BIH

EMORY UNIVERSITY SCHOOL OF MEDICINE
THOMAS K. GLENN MEMORIAL BUILDING
69 Butler Street, S.E. Atlanta, Georgia 30303

DEPARTMENT OF MEDICINE

May 21, 1981

Eugene Braunwald, M.D.
Beth Israel Hospital
330 Brookline Avenue
Boston, MA 02215

Dear Gene:

RE: JOHN R. DARSEE, M.D.

It is with great pleasure that I recommend to you Dr. John R. Darsee for promotion to Assistant Professor of Medicine in the CFT track. I have known John since the beginning of his House Staff training at Grady Memorial Hospital and the Emory University School of Medicine and have watched his clinical and academic growth through his training, which included a fellowship in Cardiology and the Chief Residency in Medicine.

John is a superior physician, with great care and concern for patients, with extreme clinical competence, with an amazingly broad range of interests and an inquiring mind, and with particular research capabilities.

He has the skill to define a question, to search for the background information necessary to answer it, and to precisely define the manner in which it is to be approached. He is extremely hard-working, but it is obvious that he enjoys his work. Coupled with this, he is an excellent teacher, and it is this feature that was most prominent during his House Staff training and particularly during his time as Chief Resident. His communication skills are superior, both in oral presentation and in writing.

As is apparent from his Curriculum Vitae, he has asked appropriate questions and answered them with competence, reflected by his multiple presentations at peer review national meetings and his publications in recognized peer review journals.

Lest this letter make him appear as an unbelievable workaholic, it deserves emphasis that he enjoys an enviable relationship with his patients and with his peers, has a dry but incisive sense of humor, and in general is an extremely well-rounded human being.

I recommend him to you unreservedly.

Sincerely,

Nanette K. Wenger, M.D.
Professor of Medicine (Cardiology)
Emory University School of Medicine
Director - Cardiac Clinics
Grady Memorial Hospital

RECEIVED
MAY 20 1981
DEPT. OF MEDICINE, BIH

NKW/jz

EMORY UNIVERSITY SCHOOL OF MEDICINE
WOODRUFF MEDICAL CENTER ADMINISTRATION BUILDING
1440 Clifton Road, N.E. Atlanta, Georgia 30322

EXECUTIVE ASSOCIATE DEAN (404) 320-5653

May 22, 1981

Eugene Braunwald, M.D.
Hersey Professor of the Theory and Practice of Physic
Harvard Medical School
Beth Israel Hospital
330 Brookline Avenue
Boston, MA 02215

Dear Dr. Braunwald:

It is a pleasure to recommend the promotion of Dr. John R. Darsee to the position of Assistant Professor of Medicine in the Harvard Medical School.

I have known John for the past six years during his post-graduate training at the Emory University School of Medicine and have continued to remain in close contact during his fellowship at the Harvard institutions. My relationship with John has been that of teacher, research collaborator, and personal friend. My understanding and admiration of his intelligence, curiosity and enthusiasm in the research laboratory, prodigious work capacity and moral character have grown steadily with time. I believe that John's accomplishments in clinical and laboratory research, his teaching and writing, and his academic leadership ability are documented in his curriculum vitae and readily apparent to all who know him.

John Darsee's academic development and contributions will surely continue on the present course and I feel confident that his association with Harvard is a rewarding experience for both parties.

Sincerely,

Don Nutter

Donald O. Nutter, M.D.
Executive Associate Dean

DON:ls

RECEIVED
MAY 26 1981
DEPT. OF MEDICINE, BIH

EMORY UNIVERSITY CLINIC

SECTION OF INTERNAL MEDICINE

1365 Clifton Road, N.E. Atlanta, Georgia 30322

(404) 321-0111

General Medicine
Paul W. Seavey, M.D.
S. Carter Davis, M.D.
Clifford F. Burnett III, M.D.
Donald C. Davis, M.D.
Bernard L. Hallman, M.D.
Larry R. Kirkland, M.D.
H. Kenneth Walker, M.D.

Gastroenterology
J. Spalding Schroder, M.D.
John T. Galambos, M.D.
Theodore Herah, M.D.
W. Scott Brooks, Jr., M.D.
Salah Nasrallah, M.D.
Stan P. Riepe, M.D.
Michael S. Perkel, M.D.

Endocrinology
E. Garland Herndon, Jr., M.D.
James H. Christy, M.D.
John K. Davidson, M.D. (Diabetes)
Paul C. Davidson, M.D. (Diabetes)
Harry K. Detcher, M.D. (Diabetes)
John R. K. Preedy, M.D.

Infectious Disease
Stephen W. Schwarzmann, M.D.
Thomas F. Sellers, M.D.
Carl Perlino, M.D.

Cardiology
J. Willis Hurst, M.D.
Edward R. Dorney, M.D.
B. Woodfin Cobbs, Jr., M.D.
Paul H. Robinson, M.D.
Stephen D. Clements, Jr., M.D.
Paul F. Walter, M.D.
John H. Stone, M.D.
Robert C. Schlant, M.D.
Spencer B. King III, M.D.
John S. Douglas, Jr., M.D.
Robert H. French, M.D.
Daniel Arensberg, M.D.
Andreas R. Gruentzig, M.D.

Hematology
Chas. C. Corley, Jr., M.D.
James W. Keller, M.D.
Chas. M. Huguley, Jr., M.D.
W. Ralph Vogler, M.D.
L. T. Heffner, Jr., M.D.
Elbert F. Winton, M.D.

Rheumatology
C. H. Wilson, Jr., M.D.
John A. Goldman, M.D.
Stephen B. Milder, M.D.

Pulmonary Diseases
Alan L. Plummer, M.D.
G. Michael Duffell, M.D.
Jeffrey R. Pine, M.D.

June 9, 1981

Eugene Braunwald, M.D.
Herrman Ludwig Blumgart Professor
of Medicine
Hersey Professor of the Theory and
Practice of Physic
Harvard Medical School

Dear Dr. Braunwald:

When I wrote you several years ago in support of Dr. John Darsee's application for a fellowship in cardiovascular research, I said that he was extra-ordinary. The past two years have only reinforced that view.

Dr. Darsee is clearly one of the most remarkable young men in American medicine. There is little need to dwell on his accomplishments in cardiovascular research. His curriculum vitae offers abundant evidence that he does original work, and that he does it in great abundance. Dr. Darsee has achieved a national reputation within a few short years, a fact that places him in a rather elite group.

I am equally impressed by Dr. Darsee's performance outside the research laboratory. He has demonstrated excellent clinical skills and teaching ability. It is not extravagant to say that Dr. Darsee became a legendary figure during his year as Chief Resident in Medicine at Grady Memorial Hospital. During that year, he did both clinical and basic research, carried out all the demands of being a Chief Resident, and took time to listen to the griefs of the medical house staff. I know less of Dr. Darsee's administrative abilities. I only had the opportunity to observe them during his year as Chief Medical Resident. He carried out those duties admirably, and had unusually good rapport with the medical house staff and faculty.

RECEIVED
JUN 15 1981
DEPT. OF MEDICINE, BIH

John Darsee, M.D.
Page Two

Dr. Darsee is a gifted, personable and incredibly well disciplined physician. His professional and personal qualifications strongly commend him for promotion.

Sincerely,

Paul Walter

Paul F. Walter, M.D.
PFW:ps

APPENDIX D

BRIGHAM AND WOMEN'S HOSPITAL - HARVARD MEDICAL SCHOOL - BETH ISRAEL HOSPITAL

EUGENE BRAUNWALD, M.D.
Hersey Professor of the Theory and Practice of Medicine

Chairman, Department of Medicine
Physician-in-Chief
Brigham and Beth Israel Hospitals

July 19, 1982

Dr. Thomas P. Stossel
Editor
The Journal of Clinical Investigation
Editorial Office
Hematology-Oncology Unit
Massachusetts General Hospital
Boston, Masachusetts 02114

Dear Dr. Stossel:

Enclosed is a letter which we are submitting for publication concerning the paper by Darsee, J. R., Kloner, R. A., and Braunwald, E.: Early recovery of regional performance in salvaged ischemic myocardium following coronary artery occlusion in the dog. J. Clin. Invest. 68: 225-239, 1981. We respectfully request the enclosed letter be published in the Journal of Clinical Investigation at the earliest possible date so that its contents will be available to the readers.

We wish also to advise you that we have made an effort to draft this letter with precision so that the scientific community will be fully informed of our opinion as to the status of this work and so that all individuals involved in this matter are treated as fairly as possible. We request therefore that the enclosed letter be printed as drafted and that its contents not be released in a form that is altered or less than complete.

The concerns expressed in the letter are in accord with discussions held with a Special Panel of the National Heart, Lung and Blood Institute and other members of the scientific community with whom we have consulted. Needless to say we regret very much that circumstances require us to take this action.

Thank you for your prompt consideration of this request.

Sincerely yours,

Eugene Braunwald

Eugene Braunwald, M.D.

Robert A Kloner MD PhD

Robert A. Kloner, M.D., Ph.D.

Enclosure

Reply to: Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115 (617) 732-6340

BRIGHAM AND WOMEN'S HOSPITAL - HARVARD MEDICAL SCHOOL - BETH ISRAEL HOSPITAL

EUGENE BRAUNWALD, M.D.
Hersey Professor of the Theory and Practice of Medicine

Chairman, Department of Medicine
Physician-in-Chief
Brigham and Beth Israel Hospitals

July 19, 1982

Dr. Thomas P. Stossel
Editor
The Journal of Clinical Investigation
Editorial Office
Hematology-Oncology Unit
Massachusetts General Hospital
Boston, Massachusetts 02114

Dear Dr. Stossel:

During 1981 we became aware of irregularities in data collected in our laboratory, as described in Section II of the Report of an Ad Hoc Advisory Committee to the Dean of the Harvard Medical School (Harvard University Gazette, January 29, 1982, Vol. 77, p. 1 and Nature, Dec. 24, 1981, Vol. 294, p. 684). This problem and further investigation have brought into question the validity of the data in the paper:

Darsee, J. R., Kloner, R. A. and Braunwald, E.: Early recovery of regional performance in salvaged ischemic myocardium following coronary artery occlusion in the dog. J. Clin. Invest. 68: 225-239, 1981.

Because of the circumstances referred to above, it is our opinion that this paper should be retracted.

Sincerely yours,

Eugene Braunwald

Eugene Braunwald, M.D.

Robert A Kloner

Robert A. Kloner, M.D., Ph.D.
Department of Medicine
Harvard Medical School
Brigham and Women's Hospital
Boston, Massachusetts 02115

Reply to: Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115 (617) 732-6340

BRIGHAM AND WOMEN'S HOSPITAL - HARVARD MEDICAL SCHOOL - BETH ISRAEL HOSPITAL

EUGENE BRAUNWALD, M.D.
Hersey Professor of the Theory and Practice of Medicine

Chairman, Department of Medicine
Physician-in-Chief
Brigham and Beth Israel Hospitals

July 19, 1982

Dr. Bernard Forscher
Managing Editor
Proceedings,
National Academy of Sciences
2101 Constitution Avenue N. W.
Washington, D. C. 20418

Dear Dr. Forscher:

Enclosed is a letter which we are submitting for publication concerning the paper by Kloner, R. A., DeBoer, L.W.V., Darsee, J. R., Ingwall, J. S. and Braunwald, E.: Recovery from prolonged abnormalities of canine myocardium salvaged from ischemic necrosis by coronary reperfusion. Proc. Nat'l Acad. Sci. 78: 7152-7156, 1981. We respectfully request the enclosed letter be published in the Proceedings of the National Academy of Sciences at the earliest possible date so that its contents will be available to the readers.

We wish also to advise you that we have made an effort to draft this letter with precision so that the scientific community will be fully informed of our opinion as to the status of this work and so that all individuals involved in this matter are treated as fairly as possible. We request therefore that the enclosed letter be printed as drafted and that its contents not be released in a form that is altered or less than complete.

The concerns expressed in the letter are in accord with discussions held with a Special Panel of the National Heart, Lung and Blood Institute and other members of the scientific community with whom we have consulted. Needless to say we regret very much that circumstances require us to take this action.

Thank you for your prompt consideration of this request.

Sincerely yours,

Eugene Braunwald

Eugene Braunwald, M.D.

Robert A Kloner no re

Robert A. Kloner, M. D., Ph.D.

Enclosure

Reply to: Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115 (617) 732-6340

BRIGHAM AND WOMEN'S HOSPITAL - HARVARD MEDICAL SCHOOL - BETH ISRAEL HOSPITAL

EUGENE BRAUNWALD, M.D.
Hersey Professor of the Theory and Practice of Medicine

Chairman, Department of Medicine
Physician-in-Chief
Brigham and Beth Israel Hospitals

July 19, 1982

Dr. Bernard Forscher
Managing Editor
Proceedings,
National Academy of Sciences
2101 Constitution Avenue N. W.
Washington, D. C. 20418

Dear Dr. Forscher:

During 1981 we became aware of irregularities in data collected in our laboratory, as described in Section II of the Report of an Ad Hoc Advisory Committee to the Dean of the Harvard Medical School (Harvard University Gazette, January 29, 1982, Vol. 77, p. 1 and Nature, Dec. 24, 1981, Vol. 294, p. 684). This problem and further investigation have brought into question the validity of the data on regional myocardial blood flow and some of the hemodynamics in the following paper:

Kloner, R. A., DeBoer, L.W.V., Darsee, J. R., Ingwall, J. S. and Braunwald, E.: Recovery from prolonged abnormalities of canine myocardium salvaged from ischemic necrosis by coronary reperfusion. Proc. Nat'l Acad. Sci. 78: 7152-7156, 1981.

Because of the circumstances referred to above, it is our opinion that the portions of this paper dealing with these data should be retracted.

However, the validity of the ultrastructural data, and of the biochemical data reported in the above listed paper has not been brought into question since the above-mentioned irregularities do not relate to such data. In addition, we have carried out additional experiments to examine the validity of the conclusions in this paper and have found that the time course of recovery of function is slow in tissue salvaged from necrosis by reperfusion. These findings have been submitted for publication.

Sincerely yours,

Eugene Braunwald

Eugene Braunwald, M.D.

Robert A Kloner MD PhD

Robert A. Kloner, M.D., Ph.D.
Department of Medicine
Harvard Medical School
Brigham and Women's Hospital
Boston, Massachusetts 02115

Reply to: Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115 (617) 732-6340

BRIGHAM AND WOMEN'S HOSPITAL - HARVARD MEDICAL SCHOOL - BETH ISRAEL HOSPITAL

EUGENE BRAUNWALD, M.D.
Hersey Professor of the Theory and Practice of Medicine

Chairman, Department of Medicine
Physician-in-Chief
Brigham and Beth Israel Hospitals

July 19, 1982

Professor L. E. Bottiger, M.D.
Chief Editor
Acta Medica Scandinavica
Kungsgatan 54
S-111 35 Stockholm, SWEDEN

Dear Professor Bottiger:

Enclosed is a letter which we are submitting for publication concerning the paper by Braunwald, E., DeBoer, L.W.V., Glogar, D. H., Darsee, J. R., Ertl, G., Ingwall, J. S., and Kloner, R. A.: Observations on the experimental reduction of infarct size and on the delayed consequences of transient myocardial ischemia. Acta Med. Scand. 210: Supplementum No. 651, 123-132, 1981. We respectfully request the enclosed letter be published in Acta Medica Scandinavica at the earliest possible date so that its contents will be available to the readers.

We wish also to advise you that we have made an effort to draft this letter with precision so that the scientific community will be fully informed of our opinion as to the status of this work and so that all individuals involved in this matter are treated as fairly as possible. We request therefore that the enclosed letter be printed as drafted and that its contents not be released in a form that is altered or less than complete.

The concerns expressed in the letter are in accord with discussions held with a Special Panel of the National Heart, Lung and Blood Institute and other members of the scientific community with whom we have consulted. Needless to say we regret very much that circumstances require us to take this action.

Thank you for your prompt consideration of this request.

Sincerely yours,

Eugene Braunwald

Eugene Braunwald, M.D.

Robert A. Kloner MD, PhD

Robert A. Kloner, M.D., Ph.D.

Enclosure

Reply to: Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115 (617) 732-6340

BRIGHAM AND WOMEN'S HOSPITAL - HARVARD MEDICAL SCHOOL - BETH ISRAEL HOSPITAL

EUGENE BRAUNWALD, M.D.
Hersey Professor of the Theory and Practice of Medicine

Chairman, Department of Medicine
Physician-in-Chief
Brigham and Beth Israel Hospitals

July 19, 1982

Professor L. E. Bottiger, M.D.
Chief Editor
Acta Medica Scandinavica
Kungsgatan 54
S-111 35 Stockholm, SWEDEN

Dear Professor Bottiger:

During 1981 we became aware of irregularities in data collected in our laboratory, as described in Section II of the Report of an Ad Hoc Advisory Committee to the Dean of the Harvard Medical School (Harvard University Gazette, January 29, 1982, Vol. 77, p. 1 and Nature, Dec. 24, 1981, Vol. 284, p. 684). This problem and further investigation have brought into question the validity of the data concerning flurbiprofen and that shown in figures 4 and 5 in the following paper:

Braunwald, E., DeBoer, L.W.V., Glogar, D. H., Darsee, J. R., Ertl, G., Ingwall, J. S., and Kloner, R. A.: Observations on the experimental reduction of infarct size and on the delayed consequences of transient myocardial ischemia. Acta Med. Scand. 210: Supplementum No. 651, 123-132, 1981.

Because of the circumstances referred to above, it is our opinion that the portions of the paper dealing with these data should be retracted.

However, the data and conclusions concerning observations on verapamil, fluorocarbons, amiodarone and converting enzyme inhibitor have not been brought into question since the above-mentioned irregularities do not relate to such data and conclusions. The validity of the ultrastructural and biochemical data related to the recovery of ischemic myocardium also has not been brought into question since the above-mentioned irregularities do not relate to such data.

In addition, we have carried out additional experiments to examine the validity of the conclusions derived from Figures 4 and 5. Insofar as the observations on flurbiprofen are concerned (Fig. 4) the conclusion indicating that drug therapy may result in both a lateral and epicardial zone of salvage has been noted in repeat studies in our laboratory using other drugs, and these have been published. Insofar as the observations

Reply to: Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115 (617) 732-6340

on delayed recovery of transiently ischemic myocardium is concerned (Fig. 5) the major conclusion has been validated in recent experiments which have been submitted for publication.

Sincerely yours,

Eugene Braunwald, M.D.

Robert A. Kloner MD PhD

Robert A. Kloner, M.D., Ph.D.
Department of Medicine
Harvard Medical School
Brigham and Women's Hospital
Boston, Massachusetts 02115

BRIGHAM AND WOMEN'S HOSPITAL - HARVARD MEDICAL SCHOOL - BETH ISRAEL HOSPITAL

EUGENE BRAUNWALD, M.D.
Hersey Professor of the Theory and Practice of Medicine

Chairman, Department of Medicine
Physician-in-Chief
Brigham and Beth Israel Hospitals

July 19, 1982

Ernest Page, M.D.
Editor,
The American Journal of Physiology: Heart
BH Box 137
The Pritzker School of Medicine
Department of Medicine
950 East 59th Street
Chicago, Illinois 60637

Dear Dr. Page:

Enclosed is a letter which we are submitting for publication concerning the paper by Darsee, J. R., Kloner, R. A. and Braunwald, E.: Time course of regional function after coronary occlusions of 1 to 120 minute-duration. Am. J. Physiol. 240: H399-407, 1981 and the paper by Kloner, R. A., DeBoer, L.W.V., Darsee, J. R., Ingwall, J. S., Hale, S., Tumas, J. and Braunwald, E.: Prolonged abnormalities of myocardium salvaged by reperfusion. Am. J. Physiol. 241: H591-599,1981. We respectfully request the enclosed letter be published in the American Journal of Physiology: Heart at the earliest possible date so that its contents will be available to the readers.

We wish also to advise you that we have made an effort to draft this letter with precision so that the scientific community will be fully informed of our opinion as to the status of this work and so that all individuals involved in this matter are treated as fairly as possible. We request therefore that the enclosed letter be printed as drafted and that its contents not be released in a form that is altered or less than complete.

The concerns expressed in the letter are in accord with discussions held with a Special Panel of the National Heart, Lung and Blood Institute and other members of the scientific community with whom we have consulted. Needless to say we regret very much that circumstances require us to take this action.

Thank you for your prompt consideration of this request.

Sincerely yours,

Eugene Braunwald

Eugene Braunwald, M.D.

Robert A Kloner MD PhD

Robert A. Kloner, M.D., Ph.D.

Enclosure

Reply to: Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115 (617) 732-6340

BRIGHAM AND WOMEN'S HOSPITAL - HARVARD MEDICAL SCHOOL - BETH ISRAEL HOSPITAL

EUGENE BRAUNWALD, M.D.
Hersey Professor of the Theory and Practice of Medicine

Chairman, Department of Medicine
Physician-in-Chief
Brigham and Beth Israel Hospitals

July 19, 1982

Ernest Page, M.D.
Editor,
The American Journal of Physiology: Heart
BH Box 137
The Pritzker School of Medicine
Department of Medicine
950 East 59th Street
Chicago, Illinois 60637

Dear Dr. Page:

During 1981 we became aware of irregularities in data collected in our laboratory, as described in Section II of the Report of an Ad Hoc Advisory Committe to the Dean of the Harvard Medical School (Harvard University Gazette, January 29, 1982, Vol. 77, p. 1 and Nature, Dec. 24 1981, Vol. 294, p. 684). This problem and further investigation have brought into question the validity of some of the data in the two papers listed below (1, 2).

1) Darsee, J. R., Kloner, R. A. and Braunwald, E.: Time course of regional function after coronary occlusions of 1 to 120 minute-duration. Am. J. Physiol. 240: H399-407, 1981.

2) Kloner, R. A., DeBoer, L.W.V., Darsee, J. R., Ingwall, J. S., Hale, S., Tumas, J. and Braunwald, E.: Prolonged abnormalities of myocardium salvaged by reperfusion. Am. J. Physiol. 241: H591-599, 1981.

Because of the circumstances referred to above, it is our opinion that paper #1 and the portions of paper #2 dealing with the data on regional myocardial blood flow should be retracted.

However, the validity of the ultrastructural, biochemical and hemodynamic data reported in the second paper (2) has not been brought into question since the above-mentioned irregularities do not relate to such data.

Reply to: Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115 (617) 732-5730

In addition, we have carried out additional experiments to examine the validity of the conclusions in these papers. The major conclusion, i.e., transiently ischemic myocardium exhibits delayed recovery has been validated in a recent series of experiments which have been submitted for publication.

Sincerely yours,

Eugene Braunwald

Eugene Braunwald, M.D.

Robert A. Kloner MD PhD

Robert A. Kloner, M.D., Ph.D.
Department of Medicine
Harvard Medical School
Brigham and Women's Hospital
Boston, Massachusetts 02115

BRIGHAM AND WOMEN'S HOSPITAL - HARVARD MEDICAL SCHOOL - BETH ISRAEL HOSPITAL

EUGENE BRAUNWALD, M.D.
Hersey Professor of the Theory and Practice of Medicine

Chairman, Department of Medicine
Physician-in-Chief
Brigham and Beth Israel Hospitals

July 19, 1982

Dr. Elliot Rapaport
Editor-in-Chief
Circulation
Building 9
San Francisco General Hospital
San Francisco, California 94110

Dear Dr. Rapaport:

Enclosed is a letter concerning the paper by Darsee, J. R., Kloner, R. A. and Braunwald, E.: Demonstration of lateral and epicardial border zone salvage by flurbiprofen using an in vivo method for assessing myocardium at risk. Circulation 63: 29-35, 1981. We respectfully request this letter be published in Circulation so that its content will be available to the readers.

We wish also to advise you that we have made an effort to draft this letter with precision so that the scientific community will be fully informed of our opinion as to the status of this work and so that all individuals involvd in this matter are treated as fairly as possible. We request therefore that the enclosed letter be printed as drafted and that its contents not be released in a form that is altered or less than complete.

The concerns expressed in the letter are in accord with discussions held with a Special Panel of the National Heart, Lung and Blood Institute and other members of the scientific community with whom we have consulted. Needless to say we regret very much that circumstances require us to take this action.

Thank you for your prompt consideration of this request.

Sincerely yours,

Eugene Braunwald

Eugene Braunwald, M.D.

Robert A Kloner MDHA

Robert A. Kloner, M. D., Ph.D.

Enclosure

Reply to: Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115 (617) 732-6340

BRIGHAM AND WOMEN'S HOSPITAL - HARVARD MEDICAL SCHOOL - BETH ISRAEL HOSPITAL

EUGENE BRAUNWALD, M.D.
Hersey Professor of the Theory and Practice of Medicine

Chairman, Department of Medicine
Physician-in-Chief
Brigham and Beth Israel Hospitals

July 19, 1982

Dr. Elliot Rapaport
Editor-in-Chief
Circulation
Building 9
San Francisco General Hospital
San Francisco, California 94110

Dear Dr. Rapaport:

During 1981 we became aware of irregularities in data collected in our laboratory, as described in Section II of the Report of an Ad Hoc Advisory Committee to the Dean of the Harvard Medical School (Harvard University Gazette, January 29, 1982, Vol. 77, p. 1 and Nature, Dec. 24, 1981, Vol. 294, p. 684. This problem and further investigation have brought into question the validity of the data in the following paper:

Darsee, J. R., Kloner, R. A., Braunwald, E.: Demonstration of lateral and epicardial border zone salvage by flurbiprofen using an in vivo method for assessing myocardium at risk. Circulation 63: 29-35, 1981.

Because of the circumstances referred to above, it is our opinion that this paper should be retracted.

However, we have carried out additional experiments which have validated the major conclusion in this paper that drugs may result in both a lateral and epicardial zone of salvage. These results have been published.

Sincerely yours,

Eugene Braunwald, M.D.

Robert A. Kloner

Robert A. Kloner, M.D., Ph.D.
Department of Medicine
Harvard Medical School
Brigham and Women's Hospital
Boston, Massachusetts 02115

Reply to: Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115 (617) 732-5519

HARVARD MEDICAL SCHOOL ◆ BRIGHAM AND WOMEN'S HOSPITAL

ROBERT A. KLONER, M.D., PH.D.
Assistant Professor of Medicine
Department of Medicine

Mailing Address:
Harvard Medical School
180 Longwood Ave.
Boston, Mass. 02115
Telephone: (617) 732-1605

July 19, 1982

Dr. Alfred F. Parisi
Editor-in-Chief
The Journal of Cardiovascular Medicine
West Roxbury Veterans Medical Center
1400 V. F. W. Parkway
West Roxbury, Massachusetts

Dear Dr. Parisi:

Enclosed is a letter which I am submitting for publication concerning the paper by Kloner, R. A., Darsee, J. K., and DeBoer, L.W.V.: Reducing Infarct Size with Coronary Reperfusion. J. Cardiovasc. Med. 6: 1173-1181, 1981. I respectfully request the enclosed letter be published in the Journal of Cardiovascular Medicine at the earliest possible date so that its contents will be available to the readers.

I wish also to advise you that I have made an effort to draft this letter with precision so that the scientific community will be fully informed of my opinion as to the status of this work and so that all individuals involved in this matter are treated as fairly as possible. I request therefore that the enclosed letter be printed as drafted and that its contents not be released in a form that is altered or less than complete.

The concerns expressed in the letter are in accord with discussions held with a Special Panel of the National Heart, Lung and Blood Institute and other members of the scientific community with whom I have consulted. Needless to say I regret very much that circumstances require me to take this action.

Thank you for your prompt consideration of this request.

Sincerely yours,

Robert A Kloner MD PhD

Robert A. Kloner, M.D., Ph.D.

Enclosure

HARVARD MEDICAL SCHOOL ◆ BRIGHAM AND WOMEN'S HOSPITAL

ROBERT A. KLONER, M.D., PH.D.
Assistant Professor of Medicine
Department of Medicine

Mailing Address:
Harvard Medical School
180 Longwood Ave.
Boston, Mass. 02115
Telephone: (617) 732-1605

July 19, 1982

Dr. Alfred F. Parisi
Editor-in-Chief
The Journal of Cardiovascular Medicine
West Roxbury Veterans Medical Center
1400 V.F.W. Parkway
West Roxbury, Massachusetts

Dear Dr. Parisi:

During 1981 I became aware of irregularities in data collected in my laboratory, as described in Section II of the Report of an Ad Hoc Advisory Committee to the Dean of the Harvard Medical School (Harvard University Gazette, January 29, 1982, Vol. 77, p. 1 and Nature, Dec. 24, 1981, Vol. 294, p. 684). This problem and further investigation have brought into question the validity of the data in Fig. 2 in the following paper:

Kloner, R. A., Darsee, J. K., and DeBoer, L.W.V.: Reducing Infarct Size with Coronary Reperfusion. J. Cardiovasc. Med. 6: 1173-1181, 1981.

Because of the circumstances referred to above, it is my opinion that the portions of this paper dealing with these data should be retracted.

However, the validity of other data in this paper are not in question, since the above-mentioned irregularities do not apply to such data. In addition, we have carried out additional experiments to examine the validity of the conclusions derived from Figure 2. These have validated the concept that reperfusion results in a reduction of infarct size and that microvascular damage and foci of hemorrhage occur within zones of myocardial cells which are already irreversibly damaged. These findings have been submitted for publication.

Sincerely yours,

Robert A Kloner MD PhD

Robert A. Kloner, M.D., Ph.D.
Department of Medicine
Harvard Medical School
Brigham and Women's Hospital
Boston, Massachusetts 02115

HARVARD MEDICAL SCHOOL ◆ BRIGHAM AND WOMEN'S HOSPITAL

ROBERT A. KLONER, M.D., PH.D.
Assistant Professor of Medicine
Department of Medicine

Mailing Address:
Harvard Medical School
180 Longwood Ave.
Boston, Mass. 02115
Telephone: (617) 732-1605

July 19, 1982

Managing Editor
American Journal of Cardiology
875 3rd Avenue
New York, New York 10022

Dear Sir:

Enclosed is a letter which I am submitting for publication concerning the paper by Darsee, J. R., Kloner, R. A.: Dependency of location of salvageable myocardium on type of intervention. Am. J. Cardiol. 48: 702-710, 1981 and the paper by Darsee, J. R., Kloner, R. A.: The no reflow phenomenon: A time-limiting factor for reperfusion after coronary occlusion? Am. J. Cardiol. 46: 800-806, 1980. I respectfully request the enclosed letter be published in the American Journal of Cardiology at the earliest possible date so that its contents will be available to the readers.

I wish also to advise you that I have made an effort to draft this letter with precision so that the scientific community will be fully informed of my opinion as to the status of this work and so that all individuals involved in this matter are treated as fairly as possible. I request therefore that the enclosed letter be printed as drafted and that its contents not be released in a form that is altered or less than complete.

The concerns expressed in the letter are in accord with discussions held with a Special Panel of the National Heart, Lung and Blood Institute and other members of the scientific community with whom I have consulted. Needless to say I regret very much that circumstances require me to take this action.

Thank you for your prompt consideration of this request.

Sincerely yours,

Robert A Kloner MD PhD
Robert A. Kloner, M.D., Ph.D.

Enclosure

HARVARD MEDICAL SCHOOL ◆ BRIGHAM AND WOMEN'S HOSPITAL

ROBERT A. KLONER, M.D., PH.D.
Assistant Professor of Medicine
Department of Medicine

Mailing Address:
Harvard Medical School
180 Longwood Ave.
Boston, Mass. 02115
Telephone: (617) 732-1605

July 19, 1982

Managing Editor
American Journal of Cardiology
875 3rd Avenue
New York, New York 10022

Dear Sir:

During 1981 I became aware of irregularities in data collected in my laboratory, as described in Section II of the Report of an Ad Hoc Advisory Committee to the Dean of the Harvard Medical School (Harvard University Gazette, January 29, 1982, Vol. 77, p. 1, and Nature, Dec. 24, 1981, Vol. 294, p. 684).

This problem and further investigation have brought into question the validity of the data in the two papers listed below (1, 2).

1) Darsee, J. R., Kloner, R. A.: Dependency of location of salvageable myocardium on type of intervention. Am. J. Cardiol. 48: 702-710, 1981.

2) Darsee, J. R., Kloner, R. A.: The no reflow phenomenon: A time-limiting factor for reperfusion after coronary occlusion? Am. J. Cardiol. 46: 800-806, 1980.

Because of the circumstances referred to above, it is my opinion that these papers should be retracted.

However, additional experiments have been carried out in my laboratory to examine the validity of the conclusions in these papers. These have validated the conclusions in the first paper (1) that drugs may result in both a lateral and subepicardial zone of salvage. These findings have been published. The observation that reperfusion results mainly in subepicardial salvage and the major concept in the second paper that

zones of microvascular damage occur within zones of myocardial cells which are already irreversibly damaged, has also been confirmed, and the results have been submitted for publication.

Sincerely yours,

Robert A Kloner MD PhD

Robert A. Kloner, M.D., Ph.D.
Department of Medicine
Harvard Medical School
Brigham and Women's Hospital
Boston, Massachusetts 02115

APPENDIX E

Requested Corrections and Additions to the Section of the Report on the Chronology of Events at the Brigham and Women's Hospital

The chronology of events, "Alleged Scientific Misconduct, Summary of Events," set out in Section VI-B of the Report (pages 10-16), omits a number of important actions which indisputably were taken by us in relation to the discovery and investigation of this matter. In addition the descriptions of several events are incomplete. Listed below are the facts which should be added to provide a complete, accurate and fair chronology of these events, with the requested additions underlined.

1. Date: May 15, 1981 (page 10)

The Report states: "A Research Fellow and a technician in the laboratory expressed suspicions concerning data irregularities in some experiments included in Dr. Darsee's unpublished abstract on infarct extension/expansion." The report should state that the abstract had not been submitted to any journal.

2. Date: May 26 - 29, 1981 (page 11)

A complete statement of the events of this period should include the following:

Dr. Braunwald conducted his own investigation of the matter. On May 26, 1981, he interviewed the two Research Fellows and the technician who had conveyed their suspicions about Dr. Darsee's work to Dr. Kloner, met with Dr. Kloner and then with Dr. Darsee. He confronted Dr. Darsee with the evidence of his misconduct on the expansion versus extension abstract. Dr. Darsee admitted that he had recreated tracings of experiments he had performed for the abstract. Dr. Braunwald then

withdrew a prior offer of a Harvard Faculty appointment which had been made to Dr. Darsee and told him that his appointment as a Research Fellow at Harvard and the Brigham and Women's Hospital would be terminated. Dr. Braunwald and Dr. Kloner began an audit of all of Dr. Darsee's data and publications. Dr. Braunwald informed Dr. Daniel C. Tosteson (Dean, Harvard Medical School) orally of the events.

3. Date: June 25, 1981 (omitted from the Report)

At Dr. Braunwald's instructions, Susan W. Adam, Administrator, Department of Medicine, Brigham and Women's Hospital, sent a memorandum to the Brigham and Women's Hospital Office of Grants and Contracts requesting that Dr. Darsee's NIH Fellowship be terminated and returned to the NIH.

4. Date: October 1981 (omitted from the Report)

Dr. Kloner and Dr. Braunwald notified two prospective employers requesting references for Dr. Darsee about the incident of misconduct involving Dr. Darsee in May.

5. Date: October 6, 1981 (page 11)

The following statement should be inserted at the end of the existing statement:

Dr. Schwarz did not suggest in his telephone call to Dr. Kloner that he or anyone else thought that there were any irregularities or that anything was amiss with the results from the Brigham and Women's Laboratory.

6. Date: October 29, 1981 (page 12)

The following statement should be inserted at the beginning of the existing statement:

While reviewing the NHLBI graphs of the AMPIM data at the Brigham Laboratory, Dr. Kloner discovered that Dr. Darsee's data showed grossly different results for dogs 1-35 and dogs 36-50. Specifically, the early dogs, 1-35, showed low

variability in blood flow in the ischemic zone, whereas the later dogs showed results consistent with those from the other laboratories. On the basis of this finding, Dr. Braunwald and Dr. Kloner determined that Dr. Darsee had committed grave scientific misconduct in connection with his AMPIM blood flow data.

7. Date: November 3, 1981 (page 13)

The following two sentences should be substituted for the existing first three sentences, which are imprecise. (There should be no dispute about this issue since the Panel has Drs. Braunwald and Kloner's letter of November 3, 1982 to Dr. Schwarz.):

Dr. Braunwald and Dr. Kloner notified the NHLBI in writing that in view of Dr. Darsee's previous misconduct and their inability to verify Dr. Darsee's flow data they could not stand behind these values produced for the AMPIM study. They stated that the data should not be included in the study, and offered to repeat the study.

8. Date: November 6, 1981 (omitted from the Report)

Dr. Braunwald and Dr. Kloner withdrew all abstracts on which Dr. Darsee was senior-author and which had been accepted for presentation to the American Heart Association.

9. Date: November 17, 1981 (omitted from the Report)

At the meeting of AMPIM investigators in Dallas, Texas, Dr. Braunwald stated that none of the Brigham AMPIM data should be included in the study.

10. Date: December 3, 1981 (omitted from the Report)

Dr. Braunwald wrote to Dr. Willis Hurst and Dr. Robert Schlant at Emory University School of Medicine to confirm the information he had provided to them in Dallas at the November meetings of the Heart Association concerning the Darsee situation. Dr. Braunwald strongly urged that an audit of Dr. Darsee's work at Emory be conducted at Emory.

11. Date: December, 1981 (page 13)

The statement concerning the pacing study dog incident is imprecise and should be amended to read as follows:

A technician in the laboratory informed Dr. Kloner that in May 1981 a Research Fellow had given her an intact heart that he said he had removed from dog D-35, which Dr. Darsee had operated upon (5/14/81), thinking that Dr. Darsee might report blood flow values on that study. The technician told Dr. Kloner that in December she had found a gamma printout of RMBF labeled D-35 in one of Dr. Darsee's notebooks, even though she had in possession of the intact heart. Dr. Kloner examined the heart in question and then reviewed Dr. Darsee's notebook and found the RMBF gamma counter sheet labeled D-35. If the intact heart did in fact come from dog D-35, it most probably could not have been used for the blood flow analysis. The technician, however, had not observed the Fellow removing the heart from the dog in question.

12. Date: March 18, 1982 (page 15)

The following statement should be added after the first sentence:

Dr. Kloner met with several members of the NHLBI Panel at Hershey and reviewed Dr. Darsee's laboratory notebooks, the laboratory log book and the remaining AMPIM tissue blocks. Dr. Kloner presented the results of a detailed analysis of the entries in the randomization cards against entries in the log book at the meeting.

13. Date: April 8, 1982 (omitted from the Report)

Drs. Braunwald and Kloner informed the NHLBI Panel in writing that: 1) In the course of their continuing investigation of Dr. Darsee's research, Drs. Braunwald and Kloner had that day uncovered evidence that many of the dogs that Dr. Darsee had claimed received radioactive isotopes may not have had radioactivity in their carcasses at the time of their disposal; 2) This evidence

could provide further confirmation that Dr. Darsee's AMPIM blood flow data were invalid; and 3) Dr. Braunwald had examined several papers senior-authored by Dr. Darsee at Emory and had become deeply concerned about the validity of the data in three of these papers.

14. Date: April 27, 1982 (page 15)

The following statement should be substituted for the existing statement, which is imprecise:

Dr. Braunwald wrote the NHLBI Panel and met with the Panel in Bethesda. Among the matters reported by Dr. Braunwald to the Panel were: 1) Dr. Kloner and he had continued to investigate the absence of dogs on which Dr. Darsee reported RMBF values from the laboratory radioactive dog disposal list; 2) The technician responsible for maintaining the radioactive dog list had confirmed that he had recorded in it all radioactive dogs disposed of from the laboratory; 3) These findings strongly suggest that Dr. Darsee did not inject microspheres into many of the dogs on which he performed experiments and therefore may not have performed the blood flow measurements he had reported; and 4) Dr. Braunwald expressed serious concern about the validity of data obtained by Dr. Darsee in three published studies carried out at Emory and requested the Panel to address the validity of Dr. Darsee's work at Emory.

15. Date: June 17, 1982 (omitted from the Report):

Drs. Braunwald and Kloner reported in writing to the Panel that they had performed microscopic analyses of remaining heart tissue from AMPIM dogs and found that specimens from dogs operated on before May 22, 1981 did not reveal microspheres, but specimens from dogs operated on after May 22, 1981 did contain microspheres, providing additional evidence that Dr. Darsee did not inject radioactive microspheres into many of the dogs on which he claimed to have performed RMBF measurements. They further offered to send the specimens to NIH

for microscopic analysis and did so. They also submitted the specimens to the office of the Dean of the Harvard Medical School for "blinded" analysis by an expert pathologist.

16. Date: July 19, 1982 (omitted from the Report)

Results of the "blinded" analysis by independent Harvard Medical School independent pathologist confirmed that "no microspheres could be detected in slides of heart tissue from dogs used in AMPDM experiments (by Dr. Darsee) between February 25, 1981 and May 26, 1981. From June 5, 1981 (two weeks after the first confrontation with Dr. Darsee because of fabricated data) through July 10, 1981, there were microspheres in at least two of the three slides from each block." These results were subsequently submitted to NIH by Dean Tosteson.

APPENDIX F

Response to the Findings and Conclusions Section of the Report (pages 18-34)

1. Section VIII-B. "Handling of Data Falsification Incident of May 21, 1981" (pages 18, 19)

Summary of Report

The Report concludes that Dr. Kloner and I acted in a prompt and reasonable manner in notifying Dean Tosteson (and officials of the Brigham and Women's Hospital) of Dr. Darsee's data falsification. The Panel acknowledges that it is "sensitive to the altered perspective afforded by retrospective view" and that NIH has not issued guidelines concerning the responsibility of grantees, sponsors or contractors to report actual or alleged misconduct, but states that the NHLBI should nevertheless have been informed in May 1981 of the original incident of misconduct by Dr. Darsee and offered the opportunity to comment on Dr. Darsee's continuation in the study. The Report states that Dean Tosteson should have implemented an audit of Dr. Darsee's work by uninvolved investigators and speculates that entrusting the investigation to an internal audit "may have contributed to the delay in discovering the extent of the data falsification."

Dr. Braunwald's Response

As explained in my Memorandum to the NIH of Comments Concerning NHLBI Panel's Report and Range of Possible NIH Actions, dated September 24, 1982 ("Memorandum"), at page 27, I agree in hindsight that it would have been preferable to have notified the NIH of Dr. Darsee's single act of misconduct in May and that it would have been preferable if an audit of Dr. Darsee's work had been conducted by uninvolved investigators. However, as stated in the Memoran-

dum at pages 27-33, and in this Appendix at pages 7-9, I believe that the Report does not provide a full description of, or give sufficient recognition to, the severe sanctions we did initiate against Dr. Darsee in May 1981, and does not present a balanced statement of the complex legal, ethical and scientific considerations we faced when we made our initial decisions. As stated in the Memorandum at page 15, I also believe that the conclusion that an external audit could have resolved the situation appreciably earlier is unsupported speculation which should be reconsidered in light of the Panel's own failure to discover conclusive evidence of Dr. Darsee's wider misconduct which Dr. Kloner and I uncovered in our own investigation.

Therefore, I respectfully request that any published NIH report include a full description of the severe sanctions initiated against Dr. Darsee by me in May 1981 and a balanced and complete statement of the complex legal, ethical and scientific considerations we faced at that time as presented in the Memorandum at pages 27-33. In addition, I believe that the Report's speculation that the internal audit may have contributed to the delay in discovering the extent of Dr. Darsee's misconduct should be rejected and not included in any published NIH report.

B. Section VIII-C. "Internal Audit Conducted by Dr. Braunwald and Dr. Kloner" (pages 19, 20)

Summary of Report

The Report agrees with Dr. Braunwald's finding "that the creatine kinase data generated by Dr. Darsee in the study with Dr. Holman cannot be believed," but asserts that the internal audit conducted in 1981 was "insufficiently rigorous and not definitive." The Report states that the audit "should have concentrated on whether data reported were internally consistent and verifiable in retrospect and on whether experiments had been performed as described."

Dr. Braunwald's Response

For the reasons set forth in my Memorandum at pages 12-15, the assertion that the internal audit conducted in 1981 was "insufficiently rigorous and not definitive" is unjustified. Therefore, I request that such statements be rejected and not included in any published NIH report.

C. Section VIII-D. "The AMPIM Study at Brigham and Women's Hospital" (pages 20-23)

The Report discusses six separate areas of which five require my response:

1. "Data irregularities prior to May 22 (dogs 1-35)"

Summary of Report

The Report concludes that "in addition to the irregularities in blood flow and heart weight noted in late October, minimal variability is present in many other measurements carried out by Dr. Darsee (on AMPIM) . . . Dr. Kloner occasionally visited the room in which experiments were performed, but took an active role in the conduct of only a few experiments. Dr. Braunwald was seldom in the laboratory and had no direct supervisory role."

Dr. Braunwald's Response

I concur with the Report's statement regarding the lack of variability on many of Dr. Darsee's measurements on the AMPIM study in addition to the blood flow and heart weight irregularities we discovered in late October 1981. The Report, however, fails to state that Dr. Kloner and I described to the Panel on March 3, 1982, prior to being shown the Panel's statistical analyses, our own concerns about Dr. Darsee's AMPIM results in addition to his blood flow measurements. Dr. Kloner specifically stated to the Panel on March 3, 1981, that he had become suspicious of the blood pressure and heart rate data which Dr. Darsee had provided in the AMPIM study. Moreover, both Dr. Kloner and I had stated to the Panel in writing on February 28, 1982, prior to our first meeting with the

Panel, that we were suspicious of the reductions in infarct size which Dr. Darsee had achieved with Verapamil (46%) and Ibuprofen (52%), believing that they were much greater than would have been expected. (Previous studies in our laboratory had indicated reductions of 23% with both agents.) On March 3, 1982, Dr. Kloner and I also discussed with the Panel our suspicions that one way the massive reductions in infarct size which Dr. Darsee had recorded could have been achieved was if he had unblinded the study and then manipulated the coronary clamp to produce smaller infarcts in drug-treated animals. We had not carried out analyses regarding these hypotheses ourselves since we considered our AMPIM experiments to be totally invalid as we had advised the NHLBI, and we had instead focused our efforts on repeating published experiments. I believe that, to be fair, any published report of the NIH should state that we had raised questions about some of these matters both in writing and orally prior to being shown the NHLBI statistical analyses on March 3.

The roles of Dr. Kloner and Dr. Braunwald in the AMPIM study are dealt with in my Memorandum at pages 15-27 and 33-35. For the reasons set forth, I respectfully request that these conclusions of the Report be rejected and not included in a published report.

2. "Adherence to the protocol agreed to by all investigators in the study"

Summary of Report

The Report states that "[a] departure from the written protocol occurred in all dog experiments" in that "[t]he same person (Dr. Darsee) carried out the surgery, randomized the dogs, and did all but one of the postmortem analyses in an unblinded fashion." The Panel concludes that "[t]he ultimate responsibility for this departure from the protocol must rest with Dr. Braunwald, the principal investigator. Drs. Kloner and Darsee must also be held responsible, because as co-investigators both attended meetings of the AMPIM investigators

and should have been thoroughly familiar with the provisions of the protocol."

Dr. Braunwald's Response

In fairness, any statement of the failure of Brigham and Women's Hospital to adhere precisely to the AMPIM protocol concerning blinded postmortem analyses should include the fact that the Panel found that two of the other three laboratories involved in the study (Duke University and Duke/VA Hospital) also failed to follow the protocol rigorously in this regard. Since the extent of the deviation at Duke University, Duke VA Hospital and the Brigham and Women's Hospital was sufficient to cause the Panel to conclude that "[l]ack of rigorous blinding has resulted in a study that does not meet its original expectations," the deviation from the protocol at Brigham and Women's Hospital, whenever metioned in a published report, should be placed in the context of the deviations at two of the other three participating laboratories.

3. "Supervision of Dr. Darsee after May 22 (dogs 36-50)"

Summary of Report

The Report states that after May 22 "Dr. Kloner was more frequently in the room in which AMPIM studies were performed (by Dr. Darsee). . . Appropriate records were retained on these dogs . . . and the pattern of variability of most measurements became similar to that reported in other AMPIM laboratories." The Report repeats the conclusion stated in Section VIII-B that "allowing Dr. Darsee to continue to do AMPIM experiments after May 22 . . . with limited supervision, in continued deviation from the AMPIM protocol, and without consulting AMPIM co-investigators, was a mistake."

Dr. Braunwald's Response

As stated in the Memorandum at page 27, had we known on May 22 the full extent of Dr. Darsee's misconduct, we certainly would not have permitted

Dr. Darsee to remain in the laboratory and perform AMPIM experiments. However, the Report does not present the considerations which led to that decision, nor does it mention in this context (although it does in other sections of the Report) that there were no guidelines or channels through which the information of Dr. Darsee's misconduct could have been reviewed with AMPIM investigators on a confidential basis. The Report also fails to describe the demonstrable benefits that have accrued to both our investigation and that of the NHLBI from our ability to compare Dr. Darsee's pre-May 22, 1981 data with his results after May 22, 1981. It was, of course, the ability to make such a comparison that led to our discovery in October of his misconduct on the AMPIM data, that provided the basis for the NIH statistical analyses in January and February 1982, and ultimately that provided us with the basis for developing conclusive evidence of his misconduct by comparing the residual radioactivity in the pre- and post-May 22 remaining AMPIM heart tissue.

Therefore, I respectfully request that any discussion of our decision to allow Dr. Darsee to remain in the laboratory under closer supervision be placed in the context of the reasons for, and benefits derived from, this decision.

4. "Other Measurements in the AMPIM Study"

Summary of Report

The Report concludes that "[t]here is no evidence that any irregularities occurred in measurements made by anyone other than Dr. Darsee. The Panel believes that Dr. Kloner should have been regularly and more directly involved in the histological measurements which represented the primary endpoint of the study, but there is no evidence that the performance of the technicians was not satisfactory." The Report suggests that there is a possibility that "the 'blinding' of ibuprofen and ibuprofen vehicle was broken," and that "[i]f so, in view of Dr. Darsee's duties in the study, it is not apparent how anyone other

than he could have been in a position to accomplish this."

Dr. Braunwald's Response

Any discussion of this issue should note, as explained in my Memorandum at page 14 and in the Appendix at pages 3-4, that the suggestion that the blinding of ibuprofen and ibuprofen vehicle may have been broken by Dr. Darsee was first made by Dr. Kloner and me in our oral, separate presentations to the Panel, prior to being shown or being told of the NHLBI statistical analyses.

5. "Residual Radioactivity in AMPIM Cardiac Tissue"

Summary of Report

This section of the Report reviews the findings Dr. Kloner and I made regarding the measurement of residual radioactivity in the remaining heart tissue from AMPIM dogs operated on by Dr. Darsee. The Report correctly confirms our finding that:

> "Although Dr. Darsee submitted regional myocardial blood flow data derived from techniques utilizing radioactive materials, the Panel believes that the evidence indicates that Dr. Darsee did not inject radioactive microspheres into AMPIM dogs before May 22, 1981. . . . Drs. Braunwald and Kloner deserve full credit for initiating the measurement of residual radioactivity in these cardiac tissues. This technique led to the important discovery that there was an absence of residual radioactivity in many of the hearts in which Dr. Darsee purported to have measured myocardial blood flow by injecting radioactive microspheres."

Dr. Braunwald's Response

We appreciate that the Report credits Dr. Kloner and me with initiating the measurement of residual radioactivity in these cardiac tissues. I firmly believe that this discovery and our discovery of the absence of microspheres in the remaining heart tissue provide the strongest direct evidence of research misconduct by Dr. Darsee other than his original admission that he recreated tracings on the extension versus expansion abstract in May 1981. As described in my Memorandum at pages 12-15, our and the Panel's investigations demonstrate that it is not enough simply to commence an investigation by outside investigators

when-suspicion of fraud surfaces. Such an audit may fail to uncover proof of misconduct unless the investigators follow a plan that includes building upon every lead and searching for places where objective tests can be performed on existing specimens. The investigative process involved in making this discovery, which we believe should be included in the Report, was the following:

During the Winter and Spring of 1982 the investigation by Drs. Kloner and Braunwald into Dr. Darsee's research continued. The Cardiac Research Laboratory maintained a list of disposal of dogs which had been rendered radioactive by the injection of isotopes in order to measure blood flow. Discrepancies were discovered by Drs. Braunwald and Kloner between the radioactive dog disposal list, maintained by the laboratory technicians, and Dr. Darsee's laboratory notebooks, throwing into question whether dogs which Dr. Darsee claimed to have given radioisotopes actually received them. In building upon these discrepancies, the residual radioactivity was then determined in tissues of animals which, according to Dr. Darsee's records, had been injected with radioisotopes and which were still available. It was discovered that, in dogs experimented upon by Dr. Darsee before May 22, 1981, residual radioactivity was not present, but in dogs experimented upon by Dr. Darsee under closer supervision after that date, significant levels of radioactivity were present. Furthermore, microscopic analysis of heart tissue showed no microspheres (particles that carry the radioactivity) in specimens from animals operated upon before May 22, 1981, but did reveal microspheres in specimens obtained from animals operated upon after that date. As the investigation by Drs. Braunwald and Kloner continued through the Spring of 1982 all results were immediately shared with the members of the NHLBI Panel and appropriate officials at NIH, as well as with the office of the Dean of the Harvard Medical School. At the request of Drs. Braunwald and Kloner these offices performed

independent analyses of the residual radioactivity measurements, and the results corroborated the findings of Drs. Braunwald and Kloner. Moreover, the Harvard Medical School independently analyzed the remaining tissue for microspheres and reached conclusions similar to those of Drs. Braunwald and Kloner. (See letter of Dr. Shore to Dean Tosteson dated July 19, 1982, Harvard Medical School, submitted to NIH.)

Since one of the important aspects of the charge provided to the Panel concerned recommending steps to deal with similar occurrences in the future, and in light of the Panel's own experience with its investigation of Dr. Darsee's research, I believe that any published report should describe the investigative process that led us to make these important findings.

D. Section VIII-E. "Publications of Dr. John Darsee from Harvard" (pages 23-25)

Summary of Report

The Report concludes that the statistical analyses performed by the NHLBI call into question most of the tabular data provided by Dr. Darsee in the five publications evaluated by the NHLBI from the Harvard Laboratory in which he was the first author. The Report also describes the findings of Dr. Kloner and me relating to the absence of residual radioactivity in the few remaining tissue blocks from the experiments recorded in Dr. Darsee's published papers and states that these findings raised "further doubts about those five papers." The Report concludes that "[e]ven if appropriate primary data were available for detailed review, it would remain difficult to vouch for the accuracy and validity of the papers in view of the extensive irregularities now apparent and the possibility of alterations in protocol which could not be appreciated by any form of data review. The primary responsibility for this situation rests with Dr. Darsee, but must be shared by his co-authors." The Report also

discusses randomization of dogs, concluding that, in the papers on which Dr. Darsee was first author, "randomization procedures and some experimental methods [are described] in a manner that does not represent the actual procedures and methods in a completely accurate way." The Report concludes that "[r]esponsibility for the laboratory's use and reporting of randomization and experimental procedures ultimately resides with Drs. Kloner and Braunwald. In the studies of Dr. Darsee in which selection of retrospective control dogs was apparently not undertaken in a random manner, Dr. Darsee would appear to be accountable for any bias that might be present in the selection process."

Dr. Braunwald's Response

I agree with the Report's conclusion that, in view of the extensive irregularities now apparent in Dr. Darsee's data and the impossibility of determining by any form of data review whether he performed experiments as he described them, that Dr. Darsee's research data must be considered unreliable. When Dr. Kloner and I reached this conclusion after completing our investigation, we immediately determined that our duty to assure the integrity of the data emanating from our laboratory and our obligation to science and the scientific community required us to communicate with the journals in which Dr. Darsee's research from Harvard had been published. Accordingly, we have submitted letters to each of the journals in which papers were published in collaboration with Dr. Darsee from Harvard, requesting them to publish retractions of articles or sections of articles as to which we had questioned the validity of the data. (Appendix D.)

The Report's statement that "the primary responsibility for this situation rests with Dr. Darsee, but must be shared by his co-authors" is, in my opinion, a casual and unjustified allegation. As stated in the Memorandum at pages 35-36, we assume a measure of responsibility as co-authors. But, as the Report con-

cluded, "[n]one of the Panel's findings suggest the involvement of any individual other than Dr. Darsee in the data fabrication at the Brigham and Women's laboratory." (Page 31.) The Report also concludes that "supervisory practices have been based, appropriately, on an assumption of honesty and trust between supervisors and trainees." (Page 28.) What our experience with Dr. Darsee shows is that a researcher determined to fabricate data and mislead his supervisors will be able to do so for a period of time under the best of conditions. As in so many other ways, Dr. Darsee misled us in his reporting of randomization procedures and in the number of dogs shared between different studies.

As we stated to members of the NHLBI Panel in writing on April 27, 1982, during the time that Dr. Darsee was in our laboratory an effort was made to randomize animals in many of the experiments. In order to economize animal usage, as well as professional and technical time, control animals which had previously been randomized were sometimes used to provide data for more than a single experiment. This was done when the investigating Research Fellow was the same and the experimental conditions were similar. Fellows have performed their own control experiments; no "control data bank" ever existed in our laboratory. Obviously, it is essential that (as in any other aspect of an experiment) the investigator be scrupulously honest in selecting controls which have previously been randomized. At the time Dr. Darsee carried out the experiments which formed the basis of several of the papers he published from our laboratory, we had no reason to doubt his integrity and we saw no reason not to allow him to use his randomized controls in more than a single experiment.

I believe that the practice of utilizing controls in more than a single experiment is not uncommon and, although it should be, it is not always spelled out in published papers. We have, however, advised the Panel that in the future we shall explain the randomization procedure in greater detail in the

Methods Section of our papers. A more detailed explanation, however, would not have eliminated the bias which appears to have occurred in Dr. Darsee's selection of control dogs.

The principles for assessing responsibility for dishonest reporting regarding randomization are really no different than for Dr. Darsee's failure to inject radioactive microspheres into dogs on which he claimed to have measured radioactive myocardial blood flow, or for any of the other instances of his misconduct.

E. Section VIII-F. "Reporting of Regional Myocardial Blood flow on a Heart that was Not Analyzed" (pages 25-26)

Summary of Report

The Report concludes that the blood flow data reported by Dr. Darsee for pacing study dog D-35 are not valid. The Report states that it would have been "preferable" for Drs. Braunwald and Kloner to have informed the Harvard and NIH Committees, but acknowledges that "uncertainties about the identity of the stored heart and other aspects of the incident might have prevented absolute conclusions from being drawn at the time."

Dr. Braunwald's Response

In contrast to the situation existing in December 1981, the circumstances now known support the contention that Dr. Darsee did not perform the blood flow measurements on pacing study dog D-35. But I would emphasize, as the Panel acknowledged, that "uncertainties" about this matter precluded any firm conclusions from being drawn about it in December, and that our discovery which did not occur until May 1982 of the lack of residual radioactivity in heart D-35 provides the strongest evidence that Dr. Darsee fabricated the RMBF data in his notebooks for this dog.

F. Section VIII-G. "Notification of NIH of Non-Renewal of Dr. Darsee's Research Fellowship" (pages 26,29)

Summary of Report

The Report observes that Brigham and Women's Hospital officials complied with their legal obligations in notifying the NIH that Dr. Darsee's appointment would be terminated effective July 1, 1981. The Report also notes that there were no NIH guidelines indicating that the reasons for the termination should have been conveyed to the NIH, but concludes that "it would have been preferable nevertheless" for this to have been done.

Dr. Braunwald's Response

I agree that, knowing what we now know, it would have been preferable to have notified the NIH of Dr. Darsee's misconduct in May, 1981. The complex considerations that led me to take the actions I did in May, 1981 are discussed in my Memorandum at pages 27-33 and in my response to Section VIII-B above, and should be included in the report, if it is to be complete and unbiased.

G. Section VIII-H. "Supervision and Instruction in the Laboratory" (pages 27-29)

Summary of Report

The Report commends the laboratory's record of productivity and the skill and dedication of its personnel, with the sole exception of Dr. Darsee. At the same time the Report questions several aspects of laboratory supervision, addressed in detail of my Memorandum at pages 15-26, and concludes that: "The supervisory practices in the laboratory, while in no way responsible for Dr. Darsee's misconduct, may have contributed inadvertently to the ease with which he was able to produce fabricated data and to the subsequent difficulty in documenting the extent of the problem. . . . As a result of the unfortunate events reported since May 1981, a genuine and determined effort has been made to improve supervision in all phases of the laboratory's operation."

Dr. Braunwald's Response

I appreciate the Report's acknowledgements that: "The laboratory has had an outstanding record of productivity since its establishment by Dr. Braunwald in 1973 [sic.1972]. It has provided research training and experience for many young investigators. . . . [Dr. Braunwald's] accomplishments in discharging [his many duties] are impressive and reflect his deep commitment, imaginative and productive approach, and remarkable efficiency, organization, and self-discipline. . . . Dr. Kloner is a well trained and capable scientist . . . . He has made a genuine effort to discharge his duties in a careful manner . . . In many respects supervision in the Cardiac Research Laboratory at the Brigham and Women's Hospital has been adequate and comparable to that provided in similar laboratories at other institutions. General supervisory practices have been based, appropriately, on an assumption of honesty and trust between supervisors and trainees. With the exception of Dr. Darsee, recent trainees have retained original records in customary and acceptable fashion. They have been allowed appropriately increasing independence as their laboratory experience has progressed. Because most incoming trainees have had superior records of previous accomplishment, they have perhaps been capable of progressing more rapidly than trainees in many other laboratories."

My detailed response to the Report's comments concerning supervision with which I disagree are set out in my Memorandum at pages 15-26. I request that the Report's comments regarding supervision be rejected and not included in any published report of the NIH.

H. Section VIII-J. "Current Status of the AMPIM Study" (page 30)

Summary of Report

The Report notes that the AMPIM protocol requiring postmortem analysis to be conducted by a different person than the individual performing the animal

surgery and the protocol concerning randomization were not always utilized at Duke University, the Duke/VA Hospital or the Brigham and Women's Hospital. The Report states that not only the Brigham and Women's but two of the other three laboratories which participated in the AMPIM study failed to provide rigorous blinding for the study. The Report concludes that "[l]ack of rigorous blinding has resulted in a study that does not meet its original expectations. . ." and, therefore, "repetition of the portion of the study at the Brigham and Women's Hospital would not accomplish the original purposes of the study."

Dr. Braunwald's Response

The failure of three of the four participating laboratories to rigorously follow the blinding procedures called for in the AMPIM protocol is discussed in my Memorandum at page 35 and in this Appendix at pages 4-5. Criticism of the Brigham and Women's Laboratory in this respect should be considered and mentioned in the context of the failures of the Duke University and Duke VA Hospital laboratories to maintain rigorous blinding.

I. Section VIII-K. "Summary"

This section of the Report repeats in cursory fashion many of the comments. to which I have responded elsewhere in this Appendix and in my Memorandum. I believe that all comments from the body of the Report which are rejected by the NIH should also be rejected where they are repeated in the Report's summary.

Harvard Medical School
25 Shattuck Street
Boston, Massachusetts 02115
(617) 732-1501
Office of the Dean

September 24, 1982

William F. Raub, Ph.D.
Associate Director for
Extramural Research and Training
Department of Health and Human Services
National Institutes of Health
Bethesda, Maryland 20205

Dear Dr. Raub:

Thank you for sharing with us a copy of the National Heart, Lung and Blood Institute Panel's final report on the Darsee matter. There has been a fair rendering of the information which has been provided by the Office of the Dean of the Harvard Medical School.

The report states, and we agree, that despite a good faith effort, the Brigham and Women's Hospital and the Harvard Medical School made mistakes in handling the case. A committee of persons not associated with the Cardiac Research Laboratory should have been appointed to investigate the matter and audit Dr. Darsee's research in June, 1981. At that time, Dr. Darsee's confessed act of data fabrication should have been reported both to the administrators of his NIH Fellowship and to the NHLBI officials responsible for the AMPIM study.

Recognizing the absence of precedent or protocol in such matters, we began deliberating what procedures should guide us, beginning with a series of faculty committee meetings last spring. On the basis of our discussions, we will soon adopt a set of principles and procedures designed to deal with a variety of circumstances and situations, among them allegations or evidence of scientific dishonesty. They include a number of safeguards, such as mechanisms intended to

Dr. Raub
September 24, 1982

Page Two

assure (1) prompt and thorough investigation by disinterested persons and (2) timely notice to and consultation with officers of affiliated institutions and agencies with a legitimate interest in the subject matter of the investigation.

We would also like to respond to the possible range of NIH actions outlined in your letter of September 15.

With regard to the offer first proposed on November 3, 1981, and subsequently reiterated by Drs. Braunwald and Kloner to repeat their portion of the study, we are in substantial agreement with the views of Drs. Braunwald and Kloner contained in their letter of September 23. The utility and appropriateness of any such repetition may be cast into doubt by the findings and conclusions of the National Heart, Lung and Blood Institute Panel in Section VIII.J ("Current Status of the AMPIM Study").

With respect to the suggestion that the Panel may recommend recovery of all or part of the funds awarded for the AMPIM Study, we have discussed this possibility with officials of the Brigham & Women's Hospital, the successor to the contractor, the Peter Bent Brigham Hospital. They and we would be willing to enter into discussions on this question should the NIH so desire.

As for the third possibility, in view of the two extensive reviews of the Laboratory that have already been completed, we would be troubled by the perceived need for, but would not object to, a requirement such as you describe for a special site visit to precede future awards to the Cardiac Research Laboratory.

I would appreciate your continuing to keep us informed of your plans for the release of the report. We hope it will be possible for us to receive a complete version at least three days before it becomes generally available because we anticipate that, with the publication of the report, there may be questions from our own constituencies as well as from the press, and we would like to be prepared to respond in a timely fashion.

Thank you for your consideration.

Sincerely,

Daniel Tosteson

Daniel C. Tosteson
Dean

HARVARD MEDICAL SCHOOL ◆ BRIGHAM AND WOMEN'S HOSPITAL

ROBERT A. KLONER, M.D., PH.D.
Assistant Professor of Medicine
Department of Medicine

Mailing Address:
Harvard Medical School
150 Longwood Ave.
Boston, Mass. 02115
Telephone: (617) 732-1606

September 22, 1982

SEP 24 1982

Dr. William F. Raub
Associate Director for Extramural
Research and Training
National Institutes of Health
Building 1, Room 107
Bethesda, Maryland 20205

Dear Dr. Raub:

Thank you for the opportunity to review the report of the NHLBI Special Panel to Review Alleged Misconduct at Brigham and Women's Hospital/Harvard Medical School as well as the range of actions likely to be considered by the NIH. Enclosed is a detailed response to the report. In this letter I should like to address such range of actions to be considered by you.

In response to the panels range of action that the laboratory be denied a chance to repeat the AMPIM study, I can only reiterate what I told the NIH nearly one year ago - that the laboratory stands ready to repeat the study without cost to the NIH, if the NIH would so desire, and if these data were to be used along with the data from the other centers.

The potential action of requiring a special review of current laboratory procedures and supervision prior to making future NIH awards is a situation that appears to be addressed already by the report. On page 3 of the NHLBI report, the panel was charged, "describe and assess the relevant practices, procedures, and supervision that existed, and that now exists, in these laboratories." At two separate times (March 4, 1982 and April 12, 1982) data and data notebooks from fellows and students who were working in the laboratory since June, 1981 were made available to the NHLBI panel. This represented the equivalent of approximately 3 man-years worth of research. The NHLBI also interviewed two fellows working in the laboratory after June, 1981. The report concludes (top of page 24) that: "since May, 1981, a genuine and determined effort has been made to improve supervision in all phases of the laboratory's operation" and on page 27 "fellows are currently required to enter specific data for each experiment conducted in the general log book as well as in their personal records" and on page 28 "In terms of newly adopted or reinforced laboratory policies, changes have been instituted over the last year with respect to: log books; adoption and implementation of new

protocols, use of a standard laboratory manual; recording of data and retention of records; data review; and allocation of dog costs to specific grants or contracts".

As the Report points out in its Summary Section (pages 31-33), this laboratory has undergone "an unusually rigorous scrutiny." Moreover, the Panel has found and so states that "apart from the professional misconduct of Dr. John Darsee, the integrity and scientific capabilities of the professional and technical staff of the Cardiovascular Laboratory at the Brigham and Women's Hospital are of a high order" and further on page 33 that "since these unfortunate events appropriate changes have been implemented by Drs. Braunwald and Kloner to improve the overall operation of the laboratory."

It is also notable that the laboratory was visited by a subcommittee of the Ad Hoc Advisory Committee to the Dean of the Harvard Medical School on dishonesty in scientific research. That Panel concluded "the subcommittee visiting the laboratory and the Committee as a whole were convinced that the laboratory directed by Dr. Kloner is well administered with appropriate regard to data collection and review. The visiting subcommittee noted the data books dating back to 1977 revealed an impressive set of well maintained, annotated and stored raw data that was in good enough condition to allow re-analysis even at this time and seemed complete with protocol, mathematical analysis and finished conclusions. The subcommittee concluded that "the present problem does not appear to be at all referable to the existing cardiac research laboratory standards, policies or procedures, nor to overt pressure provided by its director or Dr. Braunwald. Dr. Robert Kloner and the Cardiac Research Laboratory have maintained an extremely effective system for data collection, analysis and storage."

The laboratory has been formally investigated, reviewed, and inspected at least twice and the NHLBI has favorably reported on the current state of laboratory procedures. In light of these two separate reviews and their findings, as set forth above, I ask the NIH to consider very carefully the advisability of requiring yet again another site visit to review "current laboratory procedures and supervision prior to making future NIH awards to the Cardiac Research Laboratory" As a practical matter the preparation for and participation in reviews of these sort take an enormous commitment of time and divert such time and energy of the staff of the laboratory from their productive research. To require additional site visits would be duplicative and the additional cost, time and expense to both institutions would be significant. Nevertheless, should the NIH deem another site visit necessary prior to future NIH awards we will be happy to cooperate and participate. In view of the findings of the previous two site visits, I hope that it will be possible for the NIH to resolve this situation without requiring yet again another site visit.

In response to the formal NIH acceptance or rejection of particular conclusions of the panel regarding laboratory procedure, protocol, supervision, professional competence, and other matters, my response to these conclusions are set forth in detail in my attached comments. I urge the NIH to adopt and reject those particular conclusions as I have set forth.

In particular, I would like to emphasize in this letter one section of the report which should be rejected concerning use of controls in more than one study (see page 9 of my response to NHLBI statement on page 23 of their report, "the panel was told ..."). This comment in the report is taken out of context and is inappropriate, and potentially harmful.

It is important in this letter to reemphasize a general criticism of the report. Under normal circumstances the level of supervision and professional competence and laboratory procedures were on a par with other comparably-sized laboratories and were in fact adequate. Because of the abnormal circumstances of research fraud by one collaborator the NHLBI report tries to impose a set of retroactive standards. I believe that it is too easy to apply retroactive standards as a way to explain a highly abnormal circumstance and that the NIH should reject such imposition by the NHLBI panel. I am not saying that there is no room for improvement in laboratory practices, procedures or supervision (I believe over the last 16 months our laboratory has improved and evidently the NHLBI panel concurs) but to apply retroactive standards to the Darsee case is inappropriate.

With regard to handling research fraud, I believe that at the time we acted appropriately given the fact that there were no guidelines which then existed. We had to balance many factors including legal factors and the rights of Dr. Darsee which are not taken into full account in the NHLBI report. With the guidelines from the Ross committee and AAMC, if a similar episode should occur, we would definitely involve the NIH at a much earlier step and would have requested an external audit earlier. It would be helpful if the NIH could develop a procedure to allow confidential reporting of such instances by the sponsored institution.

In response to the range of action concerning notification of journals, we sent in July, 1982 letters of retraction and clarification to the journals concerning Dr. Darsee's work in our laboratory. These letters are scheduled for publication within the next few months.

Sincerely yours,

Robert A. Kloner, M.D., Ph.D.

RAK:nw

MEMORANDUM

To: William F. Raub, Ph.D.
Associate Director for
Extramural Research and Training
National Institutes of Health
Bethesda, Maryland 20205

From: Robert A. Kloner, M.D., Ph.D.

Subject: Report of the NHLBI Special Panel to Review Alleged Misconduct at the Brigham and Women's Hospital/Harvard Medical School

Date: September 14, 1982

SECTION I

GENERAL STATEMENT

I thank the NIH for the opportunity to review the Report of the NHLBI Special Panel to review alleged misconduct at Brigham and Women's Hospital/Harvard Medical School.

I agree with much of the factual material on pages 4-7. I agree that in retrospect it would have been preferable to notify NHLBI and AMPIM Co-workers about our difficulty with Dr. Darsee earlier. I agree that in retrospect an external audit performed earlier would have been preferable.

My main concerns about the NHLBI Panel Report as it currently reads is that it treats the matter from a point of view of hindsight. At the time of our initial uncovering of wrongdoing neither Harvard nor the NIH had specific guidelines on how to handle instances of research fraud, and both institutions have admitted this. Many conclusions in the report, clearly, though politely, place "blame" on Harvard or Dr. Braunwald, or myself, based upon what appears to be a policy which is only now evolving. It is unfair to apply a standard of operation or conduct retroactively, based upon perfect hindsight, and it becomes potentially destructive to impose

that retroactive standard which can then be used to criticize institutions or individuals.

To support this position I quote Science Magazine, Vol. 271, July 16, 1982. William F. Raub, Ph.D. (addressee herein) is reported as saying that with regard to misconduct, there are questions that confront NIH staffer: "What to do with allegations? What data is to be gathered? To whom to report problems, and under what circumstances? What sanctions should be considered?" "These questions", he said, "until a few months ago were largely unanswered, much less addressed."

Dr. Raub is further quoted in the article, "we need to inform them (client universities and institutions) of their obligations."

The Executive Council of the Association of American Medical Colleges has only recently adopted a statement entitled, "The Maintenance of High Ethical Standards in the Conduct of Research," dated June 24, 1982. That statement concludes that "it is unrealistic...to assume that research fraud can be entirely prevented." (pg. 2) The paper recognizes that "loss of trust because of isolated instances of dishonest behavior on the part of a few researchers could cause great harm..." (pg. 1).

These statements further show that at the time of Dr. Darsee's misdeeds, no clear-cut guidelines on how to handle the problem were available. I hope that the NIH considers the fact that policies on handling research fraud are only now evolving, as it draws its final conclusions and recommendations. I sincerely hope the NIH avoids the easy danger of condemnation of the laboratory or individuals involved through the application of perfect hindsight and a code of conduct derived therefrom as it assesses and acts upon the NHLBI report.

In focusing attention on specific areas of the report, it is important to re-

emphasize another concept that is critically important when an institution, whether it be the Cardiovascular Research Laboratory at Harvard or the NIH is being "investigated." There exists a distinction between direct responsibility and a positional responsibility when wrong doing is established. Dr. Darsee's "responsibility" was direct. The responsibility of the senior scientists because of their position exists, but I would ask the NIH to be careful not to assess blame to such senior scientists or co-authors for direct responsibility for intentional misconduct by a person in a position of trust. The distinction between direct responsibility and positional responsibility should be made in the final report.

Throughout the investigation, from the time we discovered the initial falsification by Dr. Darsee through our development of the conclusive evidence against him, we have proceeded with an honest and considered approach that had to balance our obligations to the scientific community with the forgotten issue of the NHLBI Report, the rights of Dr. Darsee. We believe that our actions were proper and best addressed those competing and complex obligations at the time. Scientific fraud by one individual was discovered and every human effort has been made to correct that fraud and devise means to prevent any such future occurrence.

I respectfully request that the NIH review the current NHLBI report and make corrections consistent with the comments above and the specific comments which follow.

SECTION II

Response to the NHLBI Panel's statements concerning Supervision and Instruction in the Laboratory

I received my research training in three separate laboratories -- one at the NIH, one at Northwestern University (now relocated to Duke University) and one at Harvard University (under the supervision of Dr. Peter Maroko). I have

been involved in cardiac research since 1972. My training at Northwestern included a Ph.D. degree as well as my M.D. I am currently an Established Investigator of the American Heart Association. I am well acquainted with what standards and supervisory practices were present in those laboratories in which I have trained and tried to have similar standards and supervisory practices when running the laboratory during Dr. Darsee's tenure. Although I believe that standards and supervision have improved since Dr. Darsee's tenure, I believe that supervisory practices had been on a par with these other laboratories. I provided a four-day orientation course on laboratory procedures and techniques, ongoing day-to-day review of primary data, formal review of new proposals, formal data meetings with the fellows and technicians every two weeks, and formal meetings with the fellows and Dr. Braunwald. Manuscripts and abstracts underwent many revisions between the fellows, Dr. Braunwald and myself. These practices under normal circumstances resulted in the graduating of several fellows from the laboratory during my brief tenure who have now assumed similar positions of responsibility: Drs. Robert Rude (Assistant Professor of Medicine at Southwestern Medical Center), Dr. Edward Brown (Assistant Professor, State University of New York at Stony Brook), Dr. Lawrence DeBoer (Research Associate at Massachusetts General Hospital), and Dr. Markku Nieminen (Associate Professor of Medicine, University of Helsinki). Under normal circumstances, these standards would be considered reasonable and usual medical research supervisory practices. I believe that the NHLBI's report unfairly uses hindsight and attempts to create a set of retrospective standards in order to criticize factors other than the highly abnormal circumstances of Dr. Darsee's intentional misconduct. I should like to analyze separately those statements which I believe require modification.

NHLBI Statement, page 27, middle of paragraph: "Whereas during Dr. Maroko's last year in the laboratory all of the fellows participated in most research protocols, and thus, shared many common experiences, the new procedures encouraged each fellow to develop and pursue his own studies." "Continuing supervision and instruction occurred primarily on an informal basis."

Dr. Kloner's Response: These statements deserve modification for two reasons. 1) The first statement infers that encouraging each Fellow to develop and pursue his own studies is somehow inappropriate, or somehow led to Dr. Darsee's data fabrication. In most laboratories Fellows are, in fact, encouraged to develop their own projects and their own protocols. 2) When I worked in Dr. Maroko's laboratory in 1976 and on a separate occasion again in 1977, each Research Fellow (including myself) was working on his own protocol. The fact that this was the case prior to 1978 should be stated, and the statement should be altered since in most laboratories fellows do develop their own protocols. Supervisory practices in the laboratory from July, 1979 to May, 1981 were similar to those during my tenure as a Fellow with Dr. Maroko, and to those in several comparable laboratories which I have visited. The term "informal basis" in the second statement is incorrect. Data were reviewed at formal meetings as noted above.

NHLBI Statement, page 27: "Without regular didactic presentation of topics such as hypothesis testing, research design, randomization, adequacy of controls, data interpretation, and statistical techniques." "[review of data] did not routinely involve re-examination of individual records."

Dr. Kloner's Response: This statement is not made in the context of the fraction of other laboratories which have regular didactic presentations of these topics. I strongly recommend the NIH carefully look at both its own intramural laboratories and comparable laboratories throughout the country

and ask if these individual laboratories have their own didactic sessions on such topics. Is this the norm? When I worked at the NIH as a student and in conversations with others, I know of many, indeed most laboratories in which there were no such presentations. On the other hand, Dr. Braunwald, in his meetings with us, covered many of these topics, albeit not as a formal course. I think it is unfair for the Panel to include these phrases, implying that somehow our laboratory was remiss in presenting formal didactic presentations to the extent that the Panel would have liked, without a comparison with other laboratories.

Similarly the NHLBI report infers that a second analysis of earlier data which had already been analyzed is routine and standard practice. Under normal circumstances, if data are carefully reviewed on an ongoing basis a reanalysis is not standard practice in most laboratories.

NHLBI Statement, page 27: "Log books were not carefully maintained in the laboratory and were not reviewed by senior staff, the Fellows' individual notebooks being regarded as the primary record."

Dr. Kloner's Response: Although having Fellows adequately maintain log books was a problem from time to time, most of the Fellows in the laboratory, other than Dr. Darsee, did, in fact, enter their experiments into the log books. The statement that the log books were not reviewed by senior staff implies that this is normal practice. Log books were not a source of primary data. Does NIH know in how many laboratories throughout the country this is a practice? How often do the senior staff review log books? It is unfair to fault us for not having senior staff review all of the log books in light of the circumstances and such reference should be deleted.

NHLBI Statement, page 28: "Dr. Kloner is a well-trained and capable scientist who was assigned major day-to-day responsibility for a busy research laboratory with several research Fellows at an unusually early stage in his research career. He has made a genuine effort to discharge his duties in a careful manner, but they appear to be too numerous and varied. He has not had an opportunity to assume these responsibilities in a gradual, phased manner."

Dr. Kloner's Response: I thank the Panel for their kind words in the first line of the statement. However, the implications and the inferences made in the later portion of the statement deserve reassessment. Similar comments on this page "limited interaction with senior scientists," and comments on page 33 of the Summary -- "Dr. Kloner's lack of previous experience in supervising a busy laboratory..." are comments which infer that more senior or experienced scientists would not have had such problems. No laboratory director, no matter how senior, is prepared or is trained to deal with the events which I have recently been through. As far as inexperience contributing to the problem, I would remind the NIH that Dr. Darsee in a totally independent laboratory, headed by Dr. Leonard Holman, Professor of Radiology at Harvard University, also fabricated data in a similar manner. I believe that the NHLBI Panel investigating this incident has failed to acknowledge in its report the nature of the acts or misconduct of Dr. Darsee, and has failed to acknowledge the way in which he carried out those misdeeds and misled his co- workers." All of us involved, including the NIH, have been victims of this misconduct.

Clearly here is an instance of harm that can be done by the report without furthering the cause of scientific research. A recommendation that new procedures as now implemented and discussed would help to discover such misconduct in the future would be appropriate. An attack on me is not.

I also remind the NIH that another group of senior scientists from the Ross Committee reviewed the question of supervision in the laboratory and stated that "The present problem does not appear to be at all referable to the existing Cardiac Research Laboratory standards, policies, or procedures, not to overt pressure provided by its director or Dr. Braunwald. Dr. Robert Kloner and the Cardiac Research Laboratory have maintained an extremely effective system for data collection, analysis and storage." I believe that supervision in the laboratory was on a par with that in other laboratories of comparable size and that supervisory practices did not promote Dr. Darsee to fabricate data nor contribute to the length of time it has required to investigate and discover all his misdeeds.

NHLBI Statement, page 28, bottom of the page: At the time, the laboratory cannot be said to have performed optimally in ensuring appropriate handling, retention and reporting of primary data after their collection by all fellows."

Dr. Kloner's Response: This is a curious conclusion since the Panel members noted with one minor exception (which is debatable) that other Fellows besides Dr. Darsee had no problem with retention and reporting of primary data. [At their visit to the Cardiovascular Research Laboratory, I showed Panel members primary data from several other Fellows, including some of my own previous work.] Again the report should be amended to reflect that the problem of retention and handling of data involved only Dr. Darsee, and did not reflect that of the laboratory in general. Not making this distinction is not only unfair to me but potentially damaging to the professional lives of the several earnest and honest young scientists who either worked or are currently working in this laboratory.

NHLBI Statement, page 28, bottom of page: "The Panel was told of one published paper in which some controls from a previously published paper by the same investigator was used, with Dr. Kloner's approval, without the text so specifying."

Dr. Kloner's Response: This comment is inappropriate, out of context and without foundation in the Report and should be deleted. It is the kind of charge which can easily be taken out of context resulting in serious damage. Furthermore, this charge, which could only be harmful to Dr. Braunwald and me was never discussed with either of us directly, nor did we have the opportunity to discuss it with the Panel. An opportunity for a full and satisfactory discussion has not been made in the body of the Report. These were control studies performed by the same investigator in which there was some overlap of a set of projects. The practice of utilizing controls in more than a single study when the investigator and experimental conditions are the same is not an uncommon practice in many laboratories and the practice is not always spelled out in detail in published papers. In this case the use of some controls by the same investigator in another study did not affect the validity or conclusions of the results. Although the use was not specified in the text, the use did not bias the results of the study in any way.

NHLBI Summary Statements, page 33, top of page: "Circumstances prevailing throughout Dr. Darsee's period in the laboratory"..."helped to create an environment that may have inhibited their [misdeeds] being uncovered."..."Dr. Kloner's lack of previous experience in supervising a busy laboratory and a hurried laboratory pace."

Dr. Kloner's Response: The statement that the laboratory circumstances helped create an environment that may have inhibited uncovering of Dr. Darsee's misdeeds and the statements concerning lack of experience in supervising a busy laboratory and hurried laboratory pace are not appropriate, fair or correct. Although discussed to some extent in my response to the draft Report, they require further analysis. There exists a practical danger in such a quick summary. These summary statements are not the same conclusions made in the body of the report and hence may easily be taken out of context and granted undue weight and influence. Unless firm evidence can be found for such statements, they can cause great and unwarranted harm and serve no purpose in advancing the integrity of medical research or emphasizing the conclusions in the body of the report.

As I have already stated supervisory practices were similar to those in the three laboratories in which I have previously worked. Virtually no one is experienced in dealing with research fraud until it occurs. Dr. Darsee performed similar misdeeds in an independent laboratory staffed by physicians more senior than myself.

Any argument concerning my "lack of experience" is refuted by the following fact: I searched for and discovered Dr. Darsee's discarded strip charts which provided key evidence in his initial admission of wrongdoing in May, 1981; I was the first person to confront Dr. Darsee with evidence of wrongdoing and the first to extract an admission of guilt; it was my idea to assess residual radioactivity in tissue which provided evidence of wrongdoing that Dr. Jerome Green has considered the most powerful since the initial admission of guilt; I first pointed out the period change before and after May 22nd to the NIH; I discovered the discrepancies between dates of entries of AMPIM dogs in the log books versus dates recorded on the randomization cards and I performed a histologic analysis which essentially proves that microspheres were not present in tissue prior to May 22nd. I would hope the NIH would consider modifying

what I believe are erroneous and incomplete statements within the summary.

NHLBI Summary Statement, page 33: "As a consequence, some laboratory procedures were less than optimal in regards to items such as ensuring proper analysis and verification of raw data, appropriateness of controls and randomization techniques, and attention to accurate description of methodology in manuscripts."

Dr. Kloner's Response: The shortfalls were those of Dr. Darsee and should not be extended to apply to the laboratory in general. To not state this would be highly unfair to the honest, hardworking young scientists in the laboratory at the time of Dr. Darsee's tenure and those who are currently working in the laboratory.

SECTION III

There are a number of other specific statements made by the NHLBI Panel which deserve comment.

NHLBI Statement, October, 1981, pg. 12 (line 4) through pg. 13 (line 13)

Dr. Kloner's Response: The discussion of the events in October, 1981, at the top of page 13 concerning the uncovering of data irregularities in the AMPIM study should be expanded. The phrasing of this entire paragraph infers that it was strictly the NHLBI who discovered a problem and notified Harvard of irregularities. Upon receiving the data in October, 1981, I had several telephone conversations with Drs. Keith Reimer and Robert Jennings who were my former mentors, concerning the difference in results between laboratories. It

should be noted that discrepancies between the results from my laboratory and the other laboratories were not the only discrepancies initially noted. For example, at the VA Laboratory there was a significant reduction in infarct size/risk region with the drug Verapamil, while this effect was not present in the Johns Hopkins laboratory.

Drs. Reimer and Jennings had been analyzing their data by plotting the ratio of infarct size to risk region versus subepicardial flow in several other previous studies, although we had not used this technique previously. Upon discussing this method of analysis with Dr. Reimer in the second or third week of October, I asked him to forward me copies of the graphs which he had drawn in order to see if they might shed some light on the puzzling results which several of us had observed. I then received from Dr. Reimer, a handdrawn version of graphs of necrosis over risk region versus flow which he had made from his own analysis. Late in the afternoon on October 29, while studying this graph and the computer printout sheets, I drew circles around dogs #36-50 which were studied after May 22, 1981 (the date on which Dr. Darsee first admitted fabricating data on an unpublished study). It was at this point that I became convinced that Dr. Darsee had fabricated the flow data in the AMPIM study and after disclosing this to Dr. Braunwald, he and I decided that I would communicate this serious concern to the NIH and co-workers the very next day. While it is true that other participants in the study, including Dr. Schwarz, might have been becoming suspicious of the data, as we were, I must stress that at no point in time did the NHLBI inform us that there were serious problems of integrity of our data.

NHLBI Statement, March 18, 1982, page 15: "Because of the widespread data irregularities in the AMPIM study, a careful review of the dog log book was undertaken. When the dates of experiments on AMPIM dogs recorded in the log were compared with those on the randomization cards, many discrepancies or omissions were noted."

Dr. Kloner's Response: This paragraph is worded to imply that this analysis was solely performed and discrepancies discovered by the Panel or NHLBI staff. This is not true. I ran a detailed analysis of the entries in the randomization cards versus entries into the log book and presented this to Drs. Morgan, Schwarz and Mr. Sullivan in Hershey, Pennsylvania, March 18, 1982.

NHLBI Statement, June 17, 1982, page 10: "Drs. Kloner and Braunwald forwarded to the NIH additional cardiac tissue samples from ten non-AMPIM dogs studied by Dr. Darsee. These tissues were measured for residual radioactivity. Although isotope-based flow data had been reported (and, in some instance published) for these animals, no residual radioactivity was detected."

Dr. Kloner's Response: Prior to sending these tissues, Dr. Braunwald and I notified the NIH of our own analysis of these animals for residual radioactivity and sent this to all Panel members. The body of the report (page 23, lines 30-34) makes it clear that Dr. Braunwald and I initiated this part of the investigation, and we request that this statement also appear under summary events so that these findings are not taken out of context.

The summary of events also neglects one area of investigation which I initiated and was carried out by Dr. Braunwald and myself concerning the presence of microspheres by histology on both the AMPIM study and in dog D-35. The results of the study were forwarded to the NIH, and we believe it should be included in

the report. We believe that the histologic absence of microspheres prior to May 22nd is, in fact, an even stronger piece of evidence suggesting that Dr. Darsee did not inject microspheres for regional myocardial blood flow measurement than even the lack of residual radioactivity within tissue.

NHLBI Statement, page 20, first paragraph, Item No.6: "Not all relevant allegations by co-workers in the laboratory were adequately investigated in 1981."

Dr. Kloner's Response: This statement is inaccurate. This should be evident in my letter to Dr. Braunwald in June, 1981, in which I spelled out the other allegations of the investigators and responded in detail to each of them. At that time we investigated to the best of our ability all relevant allegations.

NHLBI Statement, page 21, 3.b. - Conclusions: "The advisability of allowing Dr. Darsee . . . was a mistake".

Dr. Kloner's Response: Much of the evidence of data fabrication by Dr. Darsee stems from statistical analysis showing differences in standard errors, variability and presence and absence of radioactive microspheres prior to and after May 22nd. These differences in data reported by Dr. Darsee prior to and after May 22, 1981, represent the vast bulk of the material analyzed by NHLBI staff prior to our meeting in March, 1982. Dr. Braunwald has outlined to you his reasons for allowing Dr. Darsee to continue on the study. Dr. Braunwald and I provided the initial data to the NHLBI showing the problems with period changes after May 22nd in our letter to Dr. Green dated January 12, 1981. Although the NHLBI Panel questions the advisability of having allowed Dr.

Darsee to continue on the AMPIM experiments after May 22nd, this was a crucial step in proving that he fabricated data. In a sense, Dr. Darsee acted as his own control prior to and after May 22nd. If Dr. Darsee had been replaced by another investigator, proof of data fabrication would have been much more difficult. It was this period change on May 22nd in flows and heart weights which we pointed out to the NHLBI which ultimately led to the discovery that the same also applied to other parameters in the AMPIM study as well as to tissue radiation levels and the histologic presence of microspheres from both AMPIM and other studies by Dr. Darsee. Hence this period change has aided us in determining the extent of data fabrication and has allowed for a clear-cut retraction of data provided by Dr. Darsee.

In retrospect, if the same problem occurred, and knowing what we know now, including recommendations by the Ross Committee as well as the AAMC report, we would certainly notify the NHLBI and other collaborators earlier. However, at the time of the initial event in May, 1981, we had no guidelines from the NIH or elsewhere; we believed we were pursuing a fair and logical course. We were not hiding the matter. Please recall that no attempt was made to conceal Dr. Darsee's initial misconduct. Dr. Braunwald had notified the Dean of the Harvard Medical School, the Dean for Academic Affairs and the President of the Brigham and Women's Hospital.

NHLBI Statement, page 22 - 5a. Findings: "Of the first 35 dogs used in the AMPIM Study, only 15 were entered in the laboratory log (Appendix 4). Of these 15 dogs, the randomization data often did not match the data of entry into the laboratory log book. The laboratory also maintained a separate log to record the disposal of dogs that had received injections of radioactive materials. With respect to this latter book, Dr. Braunwald indicated to the

Panel that "the check revealed that all Models' dogs from #36 on (the first Model's dog after May 26), were recorded on the list, but none from Nos. 1-35 were so recorded."

Dr. Kloner's Response: Although the statement is correct, the implication is that this was not our finding. This is inaccurate and the statement should be expanded. The fact that the randomization data often did not match the date of entry into the log book was pointed out by me in my meetings with Drs. Morgan and Schwarz, and I provided them with a detailed analysis that I had performed on this matter.

NHLBI Statement, page 24, b. Conclusions: "The primary responsibility for this situation rests with Dr. Darsee, but must be shared by his co-authors."

Dr. Kloner's Response: As a co-author, it is true that I share responsibility, but the statements and conclusions of the report should be expanded to distinguish between my positional responsibility and Dr. Darsee's intentional misconduct. In the normal course of events of running a laboratory, it is not customary or usual for the laboratory director to recalculate means and standard errors for all tables when he is collaborating with a 33-year old physician who is 7 years into his post-graduate training, who has published numerous scientific articles prior to entering the laboratory, who was about to become a member of the faculty, and who had not done anything to arouse the suspicions of the laboratory directors. Furthermore, it is not customary for co-authors to hand all data over to a statistician to perform the type of analysis that the NIH performed, prior to submitting a publication. At the time these papers were submitted, we had faith in Dr. Darsee and had no reason to suspect that the tables he was showing us could have been based on fabricated data. Furthermore, the tabular data were reviewed by many reviewers of several journals, none of whom picked up

anything suspicious. Therefore, although co-authors certainly share responsibility, the statement with regard to responsibility needs clarification. Although I share responsibility as a co-author because of my position in the laboratory, direct responsibility or blame for fabricated tabular data that was provided by Dr. Darsee should not be assigned to a co-author or [regarded as concealment of intentional misconduct by a person in a position of trust]. The distinction between such positional responsibility and the individual conduct of Dr. Darsee for which only Dr. Darsee is responsible should be made.

NHLBI Statement, page 25, 2. a. Findings: "In the five papers referred to above, the Methods Sections imply that the dogs were assigned to control and experimental protocols using prospective randomization procedures. (Appendix 5) In their March 25, 1982, letter to the Panel, Drs. Braunwald and Kloner acknowledge that there were numerous instances in which animals were used in more than one study. The dogs were shared among the five papers in such a way that prospective randomization could not have been used. The mean value for the necrosis/risk region for the control dogs used in a single paper was significantly less ($p<0.03$) than that for the control dogs used in multiple papers.* The Methods Section of papers using shared dogs does not represent all of the technical procedures in an accurate fashion. One paper implies that all animals were sacrified at the end of a six-hour reperfusion period; some of these animals were reported to have been used subsequently in other protocols. This would appear to be impossible."

Dr. Kloner's Response: Dr. Darsee misled Dr. Braunwald and myself in his reporting of randomization procedures and in the number of dogs shared between different studies. The Panel has concluded that the responsibility for the laboratory's use and reporting of randomization and experimental procedures ultimately

resides with Dr. Braunwald and myself. However, Dr. Darsee reported randomization and experimental procedures to us that he was using and at the time they were reported to us, we believed him. Furthermore, he prepared drafts of the papers based on the techniques and procedures we believed he was carrying out. Again, the report needs to make a distinction between the positional responsibility of Dr. Braunwald and myself versus the intentional misconduct of Dr. Darsee. If Dr. Darsee reported to me at the time that he was doing a study in which randomization was being performed by using a random number generating computer, and I saw him use this computer, I would have no reason to question whether he was actually randomizing the animals. Thus, I cannot agree with the wording and inferences drawn from the statement that "reporting of randomization and experimental procedures ultimately resides with Drs. Braunwald and Kloner" without further explanation, especially if we were misled by the randomization and experimental procedures employed by Dr. Darsee. Furthermore, although we acknowledge the use of one animal in more than one study, Dr. Darsee clearly misled us in terms of the number of experiments that he used any given dog, and this should be included in the report.

NHLBI Statement, page 28, top of page: "In terms of newly adopted or reinforced laboratory policies, changes have been instituted over the last year with respect to: log books, adoption and implementation of new protocols; use of a standard laboratory manual; recording of data and retention of records; data review, and allocation of dog costs to specific grants or contracts."

Dr. Kloner's Response: This statement should say "over the last 16 months," rather than the last year, because of the expected date of the report.

NHLBI Statement, page 30: "An audit by persons outside the involved laboratory and department could have led to an earlier resolution of these unfortunate events."

Dr. Kloner's Response: I do agree with the Panel that an external audit would have been preferable, in retrospect. However, I disagree with the concept that it may have led to an earlier resolution of these events for several reasons: (1) Much of the investigation which led to discovery of how severe the extent of Dr. Darsee's fabrication actually was emanated from investigations carried out by Dr. Braunwald and myself and supplied to the NIH. (2) In a telephone conversation, Dr. Jerome Green told me that he thought the finding of absence of residual radioactivity from tissue prior to May 22nd was the strongest, definitive proof of misconduct outside of Dr. Darsee's initial admitted fabrication in May, 1981. Let us suppose that Dr. Braunwald and I had not discovered that Dr. Darsee failed to enter into the radioactive log books many dogs which appeared on his published papers. Without this I may not have thought of analyzing residual radioactivity in tissue. Suppose the analysis of radioactivity on residual tissue had not been done. If six months later, we, or another agency, came along and discovered the residual radioactivity problem, then the Panel would have been criticized for inability to carry out a careful audit and investigation. The major point I would like to make is that the uncovering of misdeeds in this situation has taken a long period of time, with one finding leading to another over many months. In May, 1981, when confronted with the problems with Dr. Darsee we acted as swiftly as possible. We performed our audit in good faith. I disagree with the statement that an audit by persons outside the laboratory and Department would have led to an earlier resolution when despite two outside committees we discovered almost all of the wrongdoing. The statement in its

current form is also in the nature of conjecture and inappropriate in this report.

NHLBI Statement, page 31, middle of the page: "There are statistical aberrations and apparent errors in tabular data provided by Dr. Darsee in several published papers from the Brigham and Women's laboratory in which he is a co-author. These fall into serious questions with both the reported data and the conclusions derived therefrom. The papers also include misleading descriptions of randomization procedures. The 'methods' sections do not adequately describe all relevant procedures performed on each animal in each paper."

Dr. Kloner's Response: Dr. Darsee provided the details of procedures including the randomization of procedures in his studies to us. This should be included in this paragraph. At the time he performed these studies and reported procedures, we had no reason to not believe him.

NHLBI Statement, page 32, second paragraph: "Additional questionable data were identified by Drs. Kloner and Braunwald in October and November, 1981 after they were alerted by NHLBI staff and AMPIM co-investigators to possible problems."

Dr. Kloner's Response: This statement deserves clarification. An analysis of the AMPIM study was performed in our laboratory toward the end of October as the other centers were analyzing their data. We also found problems with blood flow and heart weight, but this did not arise so much from being alerted by NIH as it did to mutual discussions with co-investigators and analysis of graphs. Furthermore, in May, 1981, we did not have available to us regional flow data from the other three centers. All we had were Dr. Darsee's data. Hence, we could not compare his standard error at that time to those found in the other centers.

SECTION IV

CONCLUDING COMMENTS

Over the last sixteen months, this laboratory has continued to function and has aimed towards high standards and achievements. Several physicians and technicians are employed in this laboratory. I have enforced policies which I believe will significantly limit the possibility of such an occurrence happening again in my laboratory. While these policies may limit this possibility, no policy can absolutely prevent research fraud. The several earnest, honest and hard working young scientists who are now working in the laboratory, the technicians who are currently employed there, all ask the same questions — What will the future of the laboratory be? How will the John Darsee affair affect future funding of the laboratory? How badly will the laboratory be Marred by the general judgmental approach of the NHLBI Report, which was written at a time when the NIH's guidelines for handling research fraud were first evolving? I request the NIH assess the long-reaching impact of this report. I hope that in reaching and reporting final conclusions and recommendations, the NIH will consider the issues which I have raised.

Robert A. Kloner, M.D., Ph.D.

BRIGHAM AND WOMEN'S HOSPITAL - HARVARD MEDICAL SCHOOL - BETH ISRAEL HOSPITAL

EUGENE BRAUNWALD, M.D.
Hersey Professor of the Theory and Practice of Medicine

Chairman, Department of Medicine
Physician-in-Chief
Brigham and Beth Israel Hospitals

December 13, 1982

MEMORANDUM:

To: William F. Raub, Ph.D.
Associate Director for
Extramural Research and Training
National Institutes of Health

From: Eugene Braunwald, M.D. EBraunwald

Subject: Scientific Misconduct of John R. Darsee, M.D.

By memorandum dated September 21, 1982, I addressed the major points of the NHLBI Panel's report on the scientific misconduct of Dr. John Darsee, a Research Fellow in the Cardiac Research Laboratory of the Harvard Medical School and Brigham and Women's Hospital from 1979 to 1981. Since then, an Emory University Committee has reported substantial new evidence indicating that Dr. Darsee had established a pattern of serious scientific misconduct well before he came to Harvard. As explained in point A of the September memorandum, a thorough understanding of the causes of Dr. Darsee's misconduct and its implications for the procedures and responses of the scientific community requires examination of the whole course of its development. By agreement with Ms. Mary Miers of the NIH, I am submitting this Addendum to the September memorandum to review recently discovered evidence of Dr. Darsee's misconduct at Emory University prior to his arrival at Harvard and its bearing on the NHLBI Panel's report.

#4

Reply to: Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115 (617) 732-6340

I. Summary of Dr. Darsee's Suspected Misconduct at Emory University

Dr. Darsee came to the Cardiac Research Laboratory from Emory University, where he had served in various capacities, including Chief Resident in Medicine and Clinical Fellow in Cardiology, from 1974 to 1979. In October 1982 Dr. Neil Moran, Chairman of the committee established by Emory University to investigate Dr. Darsee's work there, reported that the Emory Committee had uncovered serious problems with Dr. Darsee's published research papers and abstracts. The Emory Committee's discoveries are consistent with my own review of some of Dr. Darsee's Emory publications and also, I am informed, with the NIH staff's preliminary review of published work by Dr. Darsee from Emory.

According to Dr. Moran, the problems discovered by the Emory Committee reviewing Dr. Darsee's work in several laboratories at Emory appear to be very serious. The integrity of six published papers, three of which involved experimentation with human subjects, is in question as are a large number of published abstracts involving research both on patients and on experimental animals. Co-authors named on several of the Emory papers probably never witnessed Dr. Darsee performing the reported experiments or saw the data he purportedly collected. Co-authors on some of the Emory abstracts may not have been aware that Dr. Darsee was listing them as such, and several individuals who received acknowledgements in the Emory papers may not even have existed either at Emory or at Harvard. One paper apparently contains a significant misstatement of the research technique employed, possibly causing the reporting of grossly misleading results. Mislabeling has been discovered on several figures in one published work. In at least one instance, investigators in the United States and elsewhere attempted, unsuccessfully, to replicate the results reported by Dr. Darsee in a published paper now in question, causing confusion and expenditure of research funds and time.

II. Events Leading to Discovery of Suspected Misconduct

In order to place in context the importance of the Emory discoveries to an evaluation of conditions at, and the response of, the Cardiac Research Laboratory, it is necessary to chronicle the events by which Dr. Darsee's conduct at Harvard/ Brigham and Women's and Emory came into question.

A. Discovery of Misconduct at Harvard/Brigham and Women's Hospital

Dr. John Darsee came to Harvard Medical School and Brigham and Women's Hospital as a Research Fellow in Medicine in July 1979. He had the highest possible recommendations from senior faculty at Emory University. His performance during the first 22 months at Harvard and the Brigham were perceived as superlative in every respect and it was proposed that he be recommended for an appointment as an Assistant Professor at the Medical School. In May 1981, Dr. Darsee was observed by fellow researchers to be fabricating research tracings for a single, unsubmitted abstract. When confronted, he admitted to this serious misdeed. Dr. Darsee insisted that he had performed the experiments in question but had fabricated tracings for review by Dr. Kloner to replace those he had originally obtained and misplaced. The following actions were taken:

1. The plans to recommend him for an Assistant Professorship were terminated immediately. His fellowships at Harvard and the Brigham and his postdoctoral fellowship supported by the NIH were terminated.

2. Deans Tosteson and Adelstein and Dr. Robert Petersdorf (then President of the Brigham and Women's Hospital) were promptly notified, verbally and in writing.

3. A detailed analysis of Dr. Darsee's research work was immediately undertaken by Dr. Robert Kloner, Director of the Cardiac Research Laboratory, and myself.

4. Pending our review of his research work, Dr. Darsee was kept on in the Laboratory in an informal capacity, under increased supervision, to complete for two and a half months some experiments and writing on which he had been working.

5. In the fall of 1981 two prospective employers were notified of Dr. Darsee's single act of scientific dishonesty.

What was known in May 1981 was that a promising researcher with a long history of achievement had committed and acknowledged a single act of misconduct. Our initial investigation developed no evidence of further misconduct and Dr. Darsee insisted that this act was a single incident. Even as we considered what actions to take, I was receiving letters from Emory University enthusiastically recommending the character and competence of Dr. Darsee.

There were essentially three kinds of responses available at the time. At one end of the spectrum, Dr. Darsee could have been issued a severe reprimand, placed under observation and continued at the University and in the Laboratory in a probationary status. At the other end of the spectrum, he could have been expelled immediately from the Laboratory and publicly exposed for his single, admitted act of misconduct. This action, we believe, would effectively have ended his academic career. It was determined after much deliberation that the extreme sanction would, at that time, be premature. The middle ground which seemed the most reasonable, since it then appeared that only a single act of misconduct was involved, consisted of notifying appropriate University and Hospital authorities, as well as potential employers who made inquiry, withdrawing Dr. Darsee's offer of a faculty appointment at Harvard and terminating his NIH Fellowship and those at Harvard and the Brigham. Dr. Darsee was allowed to remain in the Cardiac Research Laboratory to be available for the

audit of his work and for two and a half months, to complete some ongoing experiments and writing, and so that we could observe him closely to determine what additional actions were warranted. Our decision was a difficult one, but, in the absence of guidelines or regulations, it appeared to provide Dr. Darsee with due process consistent with our obligation to ensure the integrity of the research data emanating from the Laboratory. The NIH was not notified at that time because no guidelines existed regarding notification in such circumstances, there was no legal obligation to do so, and no channels or procedures existed to ensure the confidentiality of such information pending a thorough investigation.

Dr. Darsee's experimentation, carried out under closer supervision between May 26 and July 25, 1981, was impeccable. The investigation conducted by Dr. Kloner and myself revealed no other evidence of misconduct until late October 1981. Earlier that month, the results of the NIH-funded AMPIM study, which Dr. Darsee was allowed to complete after our discovery of his misconduct in May 1981, had become available simultaneously to science administrators at the NIH, investigators in the other three participating laboratories, and Dr. Kloner and myself. Dr. Kloner discovered discrepancies between pre- and post-May 22 data reported by Dr. Darsee and immediately reported these to me. We decided together that he should immediately inform the NHLBI that we could not stand behind the data. He did so on October 29, 1981, and we then offered to repeat the experiments without cost to the Government. These matters are discussed in Section E of the memorandum of September 21, 1982.

I thereafter expanded my investigation beyond the Cardiac Research Laboratory and discovered serious inconsistencies in the results reported by Dr. Darsee from his experiments in the Nuclear Medicine Laboratory of the Brigham and Women's Department of Radiology. On my recommendation, Dr. Leonard Holman

of the Department of Radiology withdrew two papers describing these results which had been accepted for publication. Although we had at that time found no evidence of misconduct regarding Dr. Darsee's published data, all abstracts on which Dr. Darsee was senior author and which had been accepted for presentation before the Scientific Sessions of the American Heart Association were, as a precaution, withdrawn in November 1981.

Early in November, Dean Tosteson and I independently concluded that a committee of faculty outside the Department of Medicine, including representatives from outside Harvard University, should be established to examine the situation and advise whether other actions should be taken and whether any measures could be implemented to reduce the likelihood of future recurrence of such events. The committee, chaired by Dr. Richard S. Ross, studied the matter during December 1981 and January 1982 and issued a public report to the Dean at the end of January. At the time the Ross Committee rendered its report, no evidence of additional misconduct had surfaced and in the absence of such evidence involving Dr. Darsee's previously published papers, there appeared to be no need, and indeed no lawful right, to initiate retraction of those papers.

During the winter of 1982 the NHLBI established a Panel to investigate the matter further, and Dr. Kloner and I continued our investigation into Dr. Darsee's research. Many of the published experiments which Dr. Darsee had conducted at the Cardiac Research Laboratory were repeated and their general conclusions were supported. At a meeting in March 1982, Dr. Kloner and I reported to the NHLBI Panel that we had developed strong suspicions of other misconduct by Dr. Darsee in the NIH-sponsored AMPIM study and, at the same meeting, the Panel presented statistical analyses which called into question some of Dr. Darsee's published tabular data and other aspects of his AMPIM data.

In April 1982, Dr. Kloner and I uncovered and reported to the NHLBI evidence that Dr. Darsee may not in fact have injected radioactive isotopes into many of the dogs for which he had reported blood flow data measured by this method. The Laboratory maintained a list of the disposal of dogs which had been rendered radioactive by the injection of isotopes in order to measure blood flow. We found discrepancies between the radioactive dog disposal list, maintained by Laboratory technicians, and Dr. Darsee's laboratory notebooks, throwing into question whether dogs which Dr. Darsee reported injecting with isotopes actually received them.

Building upon the discrepancies mentioned above, Dr. Kloner and I caused the residual radioactivity to be determined in tissue of dogs which, according to Dr. Darsee's records, had been injected with radioactive isotopes. We discovered that residual radioactivity was not present in dogs experimented upon by Dr. Darsee before May 22, 1981, the date on which his original fabrication of data was first confirmed, but that significant levels of radioactivity were present in dogs experimented upon by Dr. Darsee under closer supervision after that date. Furthermore, microscopic analyses of heart tissue showed no microspheres (particles that carry the radioactivity) in specimens from dogs operated on before May 22, 1981, but did reveal microspheres in specimens obtained from dogs operated upon after that date.

As Dr. Kloner and I continued our investigation through the spring of 1982, all results were immediately shared with the office of the Dean of the Harvard Medical School and with appropriate officials of the NHLBI. These offices provided independent corroboration of our analyses of tissue for radioactivity, and the Harvard Medical School independently confirmed the analyses for microspheres. Dr. Hassan and then Dr. Nesson of the Brigham and Women's Hospital, and Dr. Moran of Emory University, were kept apprised of the situation as

well. Dr. Darsee has admitted no misdeeds other than the single incident of data fabrication in May 1981.

The absence of radioactivity and microspheres in dogs experimented upon before Dr. Darsee was caught fabricating tracings in May 1981 indicates that Dr. Darsee did not, in fact, inject radioactive microspheres into those dogs. Other than the single admitted instance of data fabrication, these discoveries provide the strongest direct evidence of research misconduct by Dr. Darsee at the Cardiac Research Laboratory. In its report, the NHLBI Panel states that "Drs. Braunwald and Kloner deserve full credit for initiating the measurement of residual radioactivity in cardiac tissues. This technique led to the important discovery that there was an absence of residual radioactivity in many of the hearts in which Dr. Darsee purported to have measured myocardial blood flow by injecting radioactive microspheres."

These findings, confirmed by an independent laboratory at Harvard in July 1982, made it clear that Dr. Darsee had committed serious misdeeds beyond the one he admitted to in May 1981, and provided Dr. Kloner and me with proof of the extent of his misconduct and with evidence of misconduct involving his previously published work which enabled us to seek retraction of those papers, despite the fact that repeat experiments had confirmed their general conclusions (some of these confirmations are now in press). Therefore, consistent with our obligation to the scientific community and as a precautionary measure, in July 1982, we sent letters to the editors of journals requesting them to publish retractions of suspect articles or sections of articles contributed to those journals by Dr. Darsee. At the request of the NHLBI, further public comments on the matter were withheld while the NHLBI Panel conducted its investigation and prepared its report.

B. Discovery of Misconduct at Emory

On November 17, 1981, at the time that Dr. Holman of the Brigham and Women's Radiology Department withdrew, at my suggestion, suspect papers authored by Dr. Darsee at the Nuclear Medicine Laboratory, I notified Dr. Willis Hurst of Emory University of Dr. Darsee's misconduct at Harvard. On December 3, 1981, I wrote to Dr. Hurst at Emory, strongly urging that Dr. Darsee's work at Emory be reviewed. Throughout the spring of 1982, I kept appropriate persons at Emory informed of developments in the investigation of Dr. Darsee's work at Harvard. Beginning in March 1982, as the extent of Dr. Darsee's misconduct at Harvard became clearer, I reviewed several of Dr. Darsee's published papers from Emory and reported my resulting concerns about their integrity to Dr. Moran, Chairman of the Emory Committee then established to investigate Dr. Darsee's research.

Responding to the need to keep the NHLBI Panel and staff informed of evidence concerning the extent of Dr. Darsee's misconduct as it developed, I also reported these concerns to them. The Panel, however, was charged or interpreted its charge to limit its review to Dr. Darsee's misconduct at the Cardiac Research Laboratory. It therefore did not act upon my repeatedly stated concerns about the Emory work, and in its report omitted any reference to evidence of Dr. Darsee's misconduct at Emory. Moreover, it did not acknowledge that a potential problem with Dr. Darsee's work at Emory had been brought to its attention.

Between April and September 1982, I remained in contact with Dr. Moran at Emory, informing him of our findings at Harvard and continuing to stress the need for prompt action and response by Emory. In response to Dr. Raub's request for my comments on the Panel's report, I prepared, in September, a written statement describing my specific concerns with several of the Emory

papers which I had previously expressed orally to the Panel, and submitted this statement to the NIH (Appendix A to Memorandum of September 21, 1982) and to Dr. Moran at Emory. On October 12, 1982, Dr. Moran informed me by telephone that his Committee had discovered serious problems regarding Dr. Darsee's work from Emory. Dr. Moran then continued to inform me of the progress of the Emory Committee and requested specific information to aid in the Emory investigation. At Dr. Raub's request, I continued to keep the NIH informed of additional information as it developed regarding Dr. Darsee's published work from Emory. Dr. Raub subsequently related that he had been informed by Dr. Moran that the Emory Committee had uncovered serious problems regarding Dr. Darsee's Emory research, and that the NIH staff's preliminary review confirmed the existence of such problems.

III. Significance of Misconduct at Emory to NHLBI Panel's Report

The Panel's report concludes that "[a]part from the professional misconduct of Dr. John Darsee, the integrity and scientific capabilities of the professional and technical staff of the Cardiovascular Laboratory at the Brigham and Women's Hospital are of a high order." (Page 32.) In responding to the discovery of a single admitted instance of misconduct, ". . . Hospital officials complied with their legal obligations, and there were no official NIH guidelines indicating that they should have informed the NIH" of their reasons for terminating Dr. Darsee's NIH fellowship. (Page 26.) "In many respects supervision in the Cardiac Research Laboratory . . . has been adequate and comparable to that provided in similar laboratories at other institutions." (Page 28.) The report confirms that, after the April 1981 discovery, Dr. Kloner and I "began an audit of all of Dr. Darsee's data and publications from the Cardiac Research Laboratory" (page 11) and that, after the NHLBI established its own Panel in December 1981, "[a]t all times, Drs. Tosteson, Braunwald, Kloner, and all

other officials and employees at Harvard and the Brigham and Women's Hospital cooperated fully and promptly in providing all requested information and in facilitating the Panel's study in every way possible" (page 32).

The Panel also concluded, and I agree, that in retrospect it would have been a preferable course to have notified the NIH immediately of Dr. Darsee's single, admitted act of misconduct. I also agree with the Panel's conclusion that the more prudent policy in such circumstances is to have uninvolved investigators perform an audit, and this institution's guidelines now provide for this procedure. At the same time, I would emphasize that it was our decision to allow Dr. Darsee to continue working at the Laboratory under close supervision after the May discovery that led to our subsequent discovery, based on discrepancies between Dr. Darsee's pre- and post-discovery experiments, of misconduct in the AMPIM study; and it was the discovery of AMPIM misconduct that led, in turn, to the discovery of Dr. Darsee's misdeeds in several laboratories of two Universities.

I have requested respectfully that speculation appearing in the Panel's report (pages 28 and 31), regarding the effect of supervision in the Laboratory be reconsidered in light of the newly discovered evidence not considered by the Panel that Dr. Darsee had previously engaged in extensive fabrication of data and other extraordinary misconduct in published work from Emory University. Additional comments on the report's statements regarding supervision in the Laboratory are contained in Point C of the memorandum of September 21, 1982. If the Emory Committee's initial findings are confirmed, Dr. Darsee developed his pattern of scientific misconduct, including his ability to avoid detection, well before he came to Harvard. In addition, his undetected misconduct at Emory helped him to amass an impressive record of significant publications and enthusiastic professional recommendations which appeared to entitle

him to substantial responsibility and trust in subsequent positions. The fact that Emory University did not uncover evidence of Dr. Darsee's misconduct either before or promptly after notification in November 1981 of his Harvard misconduct is further evidence of the problems of detecting research misconduct. Neither my initial review, nor the Ross Committee, nor the NHLBI Panel uncovered conclusive evidence of the extent of Dr. Darsee's misconduct. Against this background, the evidence developed in April and May 1982 by Dr. Kloner and me and confirmed by independent investigation at Harvard in July, 1982 that Dr. Darsee failed to inject radioactive microspheres into dogs on which he reported radioactivity measurements, is a significant discovery in a case where the researcher denies wrongdoing.

The events at Harvard focused the attention of this institution on the need to reduce the risk of dishonesty, and the NHLBI Panel has approved specific changes already implemented in the Cardiac Research Laboratory for this purpose. But just as Dr. Darsee's case cannot be dismissed as an isolated incident, the causes of his misconduct cannot be meaningfully addressed by reviewing any isolated segment of his career. The evidence of Dr. Darsee's prior misconduct, undetected for years at other laboratories, underscores the need of the scientific community for inquiry crossing institutional boundaries and for standards of general application. Unfortunately, neither the NIH nor the Association of American Medical Colleges had provided guidelines for dealing with suspected fraud in May 1981. In recent months, however, the Harvard Medical School has developed a process, the AAMC has promulated standards and the NIH has begun to develop guidelines for handling incidents of suspected fraud. Hopefully, the costly lessons learned from this case and others like it will ultimately assist institutional officials, department chairmen and laboratory directors to prevent and respond to such problems in the future.

BRIGHAM AND WOMEN'S HOSPITAL - HARVARD MEDICAL SCHOOL - BETH ISRAEL HOSPITAL

EUGENE BRAUNWALD, M.D.
Hersey Professor of the Theory and Practice of Medicine

Chairman, Department of Medicine
Physician-in-Chief
Brigham and Beth Israel Hospitals

February 11, 1983

Dr. William F. Raub
Associate Director for Extramural
Research and Training
National Institutes of Health
Bethesda, Maryland 20205

Dear Dr. Raub:

We have received and reviewed a copy of the Decision Document of the Director of the NIH dated January 5, 1983 concerning the investigation of alleged misconduct at the Brigham and Women's Hospital, together with the several attachments. We are very appreciative of the attention and care you and your staff have devoted over many months to this matter.

We wish to bring to your attention a statement in the Decision Document which we believe to be the result of a misunderstanding. The last sentence in the Decision Document (at page 9) states that the AMPIM study was "irrevocably compromised because of the failure to inform the NHLBI and the AMPIM coinvestigators." We believe that it is indisputable that it was not the failure to inform the NHLBI and AMPIM coinvestigators of Dr. Darsee's misconduct in May 1981, that irrevocably compromised the study. Rather, the Brigham and Women's Hospital portion of the study was, unknown to us at the time, already invalid and beyond salvage by May, 1981 because of Dr. Darsee's misconduct prior to that date. Notification of the NIH in May of his single act of misconduct could not have changed that fact.

As we have stated in our responses to the Panel's report, we agree, in retrospect, that it would have been preferable to have advised the NIH of that incident. However, even if we had made such notification to the NHLBI and AMPIM coinvestigators in May, the study was already irrevocably compromised because of Dr. Darsee's misconduct on experiments performed before that date on 35 of the AMPIM dogs. Terminating the Brigham and Women's Hospital portion of the study at that point or replacing Dr. Darsee for the remaining 15 dogs could not have salvaged the Brigham's study. Only if we had repeated our entire portion of the study, could that portion have been salvaged, and repetition could have occurred promptly after October, 1981, as we offered, when Dr. Darsee's misconduct on AMPIM was first uncovered. Moreover, permitting Dr. Darsee to continue on the study after May in fact did not damage the study further, but it did allow us in October, by comparing his pre- and post-May 22 results, to develop the evidence that the study had already lost its integrity because of Dr. Darsee's misconduct prior to May 22.

Reply to: Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115 (617) 732-6340

A further relevant point is that the Brigham and Women's Laboratory was only one out of five laboratories participating in the study. The tragic invalidity of the results from the Brigham and Women's Hospital did not irrevocably compromise the data provided by the other four centers participating in the study. Also, the NIH Panel reported that for reasons entirely unrelated to the failure to notify the NIH, the whole study did not meet its original expectations. The Panel noted that blinding of the study was not carried out according to the AMPIM protocol in three of the participating centers and concluded that "[L]ack of rigorous blinding has resulted in a study that does not meet its original expectations. Thus, repetition of the portion of the study at the Brigham and Women's Hospital would not accomplish the original purpose of the study." Notification of the NIH of Dr. Darsee's misconduct in May could not have altered these facts.

We respectfully request that this statement in the Decision Document be corrected.

Sincerely yours,

Eugene Braunwald

Eugene Braunwald, M.D.

Robert A. Kloner

Robert A. Kloner, M.D., Ph.D.

APPENDIX 5

PRESS COMMENTS ON DARSEE CASE

Harvard Delays in Reporting Fraud

"Six months passed before officials made public a confessed case of fraud, and then only after NIH questioned suspicious data"

In May 1981, several young researchers at the Harvard Medical School watched in astonishment as one of their colleagues flagrantly concocted data for an experiment. Here was proof, the stunned eyewitnesses later told their superiors, of something they had suspected for months—that their colleague was systematically faking a major part of his prodigious output.

John Roland Darsee, 33, with nearly 100 published papers and abstracts to his name, confessed to that one case of fraud but denied any other improprieties. Faced with a serious problem, Harvard Medical School authorities, including dean Daniel C. Tosteson, immediately relieved Darsee of his position and then began an examination of his vast opus. The audit, conducted over the course of the summer and fall, revealed that the original data for many of Darsee's experiments at Harvard were unaccountably missing. Despite the increasing gravity of the situation, Harvard authorities did nothing to make it public. Moreover, Darsee, who had been stripped of his position but not of his association with Harvard, continued to work at the lab and publish papers as if everything were just fine.

Harvard took public action 6 months after the admitted falsification, but only when the National Institutes of Health (NIH) compelled it to account for scientific discrepancies in a $724,154 study in which Darsee played a major role. The NIH study was designed to assess whether certain drugs limit the damage done by heart attacks. In answering NIH queries, Harvard officials explained publicly for the first time that the problems had to do with Darsee and his inability to produce certain critical raw data.

Harvard then withdrew his results from the study, canceled his presentation of papers at a meeting, released a brief statement to the press about the admitted falsification, and formed a blue-ribbon committee to look into the matter and make recommendations about possible further action. A summary of the findings will be issued in February. NIH is also investigating the affair.

Even though Harvard officials began their investigation of Darsee's work in May, it was not until NIH questioned results that Harvard discovered the dearth of raw data for the NIH study. At least one faculty member has criticized Harvard for delays and lack of thoroughness. Not the least troubling question in the Darsee episode is why he was allowed to remain in the lab for 6 months, especially considering the gaps in his raw data and his colleagues' contentions that much of his work had been systematically created out of whole cloth.

Harvard asserts that everything published by Darsee has been "verified," but a detailed recounting of the story raises questions on the issue.

If some of the work with Darsee's name on it were to be expunged from the literature, as might be the recommendation of the blue-ribbon committee, it would be a blow to Darsee's mentor and chief coauthor, Eugene Braunwald, 51, holder of one of the most distinguished chairs of medicine in the nation and a member of the National Academy of Sciences.

Much of Darsee's work was aimed at a hot clinical topic—how drugs and other interventions can dramatically aid recovery from a heart attack. Until the 1960's, it was thought that when the blood supply to a portion of the heart muscle was sharply reduced, the muscle died and there was little or nothing to be done about it. However, in the past decade research has shown that there is a borderline zone that may recover, and that the zone at times can be extended with various treatments.

Darsee arrived at Harvard ready to address such problems, already much praised for his quick and elegant work. He received a medical degree from Indiana University in 1974 and then proceeded to Emory in Atlanta for various residencies, attaining a remarkable pace of publication toward the end of his stay. His papers, according to at least one with whom he worked, were nothing short of magnificent. "I think John was the outstanding person I've worked with in 13 years on the faculty here," says Donald O. Nutter, executive associate dean of the school of medicine and one of Darsee's many coauthors. "He was brilliant, intellectually brilliant. He was the hardest working individual that I've ever come in contact with."

In July 1979, Darsee joined Braunwald, the Hersey Professor of the Theory and Practice of Physic and physician-in-chief of the Peter Bent Brigham and Women's Hospitals. Darsee worked doggedly to advance his career, spending most of his waking hours in the lab. He was paid $22,000 in 1980 on an NIH fellowship in addition to his small Harvard salary. It was a pittance compared with his earning potential had he entered the private practice of medicine. What Darsee chose instead was the prestige, status, and honor of performing research in an internationally respected lab.

Eugene Braunwald came to Harvard in 1972, his vitae at the time listing more than 450 papers and scores of memberships, honors, and awards. At Harvard, Braunwald's biomedical empire expanded, and along with it, new opportunities for his students. By 1980 he had pulled off a remarkable consolidation that joined his own department at the Brigham with that of the Beth Israel Hospital.

Darsee flourished in this environment. Between 1979 and 1981 he published close to 100 papers and abstracts, many of them coauthored with his mentor. When Braunwald's two-volume, 2000-page magnum opus came out in 1980 (*Heart Disease: A Textbook of Cardiovascular Medicine*), Darsee was the coauthor with Braunwald of the chapter on diseases of the pericardium. In 1980, Darsee was made an instructor at Harvard's Department of Physiology, and a full-time appointment to the faculty seemed not far off. Best of all, Braunwald, who already presided over two complete laboratories, at least $3.3 million of NIH funds, and nearly a score of researchers, was considering setting up a separate lab for Darsee at Beth Israel. In the competitive world of Boston biomedicine, a promotion of this kind, at Darsee's early age, would ensure a dazzling career.

Such was the situation in the spring of 1981. Yet there were those who were far from enthralled by Darsee's rapid climb up the biomedical ladder and his prodigious record of publication. The people

p. 478 - 482.
0036-8075/82/0129-0478$01.00/0 Copyright © 1982 AAAS

with whom he worked daily—two fellows and a technician in the lab—were suspicious of his vast output. "This thing did not come up over night," says Edward Brown, 34, who has since gone on to a faculty position at Stony Brook. "You suspect it for months." Based on work they had seen Darsee perform, Brown and the two others deduced that an abstract Darsee was preparing to send off for publication contained not a shred of actual research. According to Robert A. Kloner, who runs the lab at the Brigham, the three came forward on 18 May with charges of fakery.

Kloner is listed as a coauthor on many of Darsee's papers, and the allegations must have initially sounded farfetched. "They came to me," he says, "with their suspicions but no proof." Kloner nonetheless investigated and found what he calls "irregularities." He then asked Darsee to show him the raw data—the mass of material that, in a cardiovascular lab, can include preserved sections of dog hearts, electrocardiogram tracings, lab notebooks, and myriad measurements of blood flow and chemical composition. Darsee said he would pull some of them together.

What he did instead, according to Kloner, was to go into a room within the laboratory complex and start taking hemodynamic readings off one of the dogs in an experiment. As the chart paper came out, he marked it day 1, day 2, and so forth, making it look as though the data had been taken over the course of several days. All the while, several of his colleagues were watching the performance in awe, not believing Darsee would ever be so brazen. When later confronted by Kloner, Darsee admitted to the falsification but said he did it only because he had thrown the real raw data away. This short retention of raw data was for an experiment allegedly performed only months earlier. The exact aim of the fudged experiment and the names of the coauthors on the abstract have never been made public. Soon after the admission of guilt, the fellows told Kloner they suspected Darsee was cheating in a systematic way in order to achieve results so consistently clean and voluminous.

How could such a situation ever have occurred? A lab chief might well be in the position of knowing first hand what work was being done in the lab, although this is not always the case. In any event, it was not Kloner's responsibility to monitor the generation of all the raw data. Part of the problem was that the lab was populated with a loose federation of aspiring young researchers—often working

Coping with Fraud

A key issue in the Harvard affair is whether the senior members of the lab were sufficiently prompt in bringing Darsee's admitted falsification to the attention of federal officials, in subjecting his vast opus to critical scrutiny, or in alerting the scientific community to potential problems. Harvard is leaving judgment on these issues to the blue ribbon committee. "We had confidence in the work that was published," says Kloner, "but that is part of the purpose of the investigation: to help us determine what additional measures are necessary." Asked if those measures might include the retraction of papers, Kloner replied: "It depends on what they feel. I think the people on the committee are high-powered and we will respect their opinion . . . There are no rule books that one can go to. That is part of the purpose of the committee."

Though rule books do not exist, other institutions have dealt with incidents of falsification in a manner quite different from Harvard. In the celebrated 1974 case of the patchwork mouse, Robert A. Good, then president of the Sloan Kettering Institute for Cancer Research and mentor of the scientist under fire, relieved William T. Summerlin of his post on the same day as his admission of fakery and proceeded to launch an investigation into all his work. "Good knew," according to a book on the affair, "that his stance must be one of rigid adherence to the code that binds all scientists together. Under this code, any misrepresentation of an experiment would jeopardize the offender's career unless and until a committee of his peers has investigated the matter and found him blameless." Less than 2 months later, a five-person committee from Sloan-Kettering recommended that Summerlin's "association with the Institute be terminated."

One of the most publicized cases occurred in 1980, when a junior researcher at Yale medical school, Vijay Soman, in the course of a scientific audit admitted to falsifying work in a single paper. One week later, he was asked to resign by his boss and coauthor, Philip Felig. All the data from his lab were immediately impounded in preparation for an audit by a researcher from outside Yale (*Science*, 3 October 1980, p. 38). When the audit revealed not only scientific problems but a dearth of raw data, papers by Soman and Felig were retracted, eventually a dozen in all. Most of these retractions were the result of questions, not proof, concerning duplicity. Raw data were available for only three of the retracted papers. The rest had come under a cloud.

A point on which Yale officials have been criticized is the delay with which they investigated the initial charges, which were brought by a young NIH researcher. It took almost a year of pressure before the first audit brought about Soman's admission of fakery. During a congressional hearing on fraud in biomedical research, held in March and April of 1981, senior researcher Felig said that part of the reason for the delay was that he had ignored the raw data for the experiment under fire. "Accepting Dr. Soman's list of patients rather than examining the patient's charts was a mistake," he told Representative Albert Gore, Jr.'s (D-Tenn.) subcommittee on investigations and oversight. "A subsequent review of the charts failed to provide substantiation of the diagnosis of anorexia nervosa in at least one of the patients and this would have provided evidence casting doubt on the veracity of the data."

The Darsee affair at Harvard in some respects shows similarity with other episodes. Harvard officials, like those at Yale, seem to have neglected to speedily examine all the raw data—otherwise the problems with the AMPIM study might have emerged earlier. In other respects, Harvard stands in contrast. It waited 6 months after an admission of fakery before calling together its committee, and during that time allowed Darsee to remain in the lab with whatever data he had in his possession, raw or otherwise.

—William J. Broad

0036-8075/82/0129-0479$01.00/0 Copyright © 1982 AAAS

with little supervision. "For Darsee's earlier work," says Braunwald, "the reason we have complete confidence is that the raw data were inspected at the most fundamental level at the time they were gathered. But then his role changed. After 18 months or so it is not the custom to hold onto a fellow's hand at every turn." Even when supervision was close, it was carried out by the youngest of Braunwald's lab chiefs. Kloner, who runs the lab at the Brigham for Braunwald, is in fact a year younger than Darsee. Kloner himself had been a fellow in the lab until he was appointed chief in July 1979.

Kloner reported the misdeed to Braunwald at the end of the week, Friday, 22 May. A brief statement about the events was put out to the news media 7 months later by Tosteson. "An intense investigation was immediately undertaken, in the course of which the individual admitted this serious misconduct. His academic and clinical appointments at Harvard and at the Brigham and Women's Hospital as well as his Fellowship from the National Institutes of Health were promptly terminated. None of his work that could not be verified has been presented to the scientific community."

This statement is a condensed and in certain ways misleading version of what happened. Harvard indeed terminated Darsee's fellowship in June, but did not tell NIH why. There was, moreover, no immediate move by his superiors at Harvard to pull work Darsee had submitted for publication or to stop him from publishing other work—even though, according to Braunwald, the examination and verification of his work was not completed until that fall. Darsee may have lost his appointments in May, but these papers and abstracts list his address as the "Brigham Hospital and the Harvard Medical School." At the end of May, for instance, ten of Darsee's abstracts were accepted by the monthly journal *Circulation*, in anticipation of his presenting papers at the annual meeting of the American Heart Association (AHA) in November. Darsee is lead author on nine of the abstracts. Of the 4000 or so authors whose abstracts appear in the October *Circulation*, Darsee is topped by only one, who has 11 abstracts. According to AHA officials, the abstracts could have been withdrawn up until the very end of July.

Six of the Darsee abstracts are from Harvard, the rest listing his address as Emory. Five of the Harvard abstracts are coauthored with Kloner and Braunwald. In addition, papers by the trio kept appearing in publications during the summer and fall. Darsee as lead author and Kloner as coauthor have a detailed, nine-page report in the October issue of the *American Journal of Cardiology*, describing drugs that limit the damage caused by heart attacks. In the November issue of the *Proceedings of the National Academy of Sciences*, a paper appears that is coauthored by Braunwald, Kloner, and Darsee. A line at the top of the article reads: "Contributed by Eugene Braunwald, August 3, 1981."

Perhaps Harvard officials thought that Darsee would eventually be rehabilitated. Certainly, Harvard took no public action during this 6-month period that would have hindered Darsee from resuming his career.

Stripped of his appointments but not of his association with Harvard, Darsee remained in the lab. In particular he was allowed to work on a $724,154 multi-institutional study sponsored by NIH. This study was known as AMPIM (Animal Models of Protecting Ischemic Myocardium), in which drugs were tested on dogs to see if they could limit the damage done by heart attacks. In addition to Braunwald's lab at the Brigham, groups were participating from Duke, the Duke/VA Medical Center, and Johns Hopkins. The study had been under way at Harvard since September 1979, and Harvard officials decided to let Darsee continue to work on the project throughout the summer. Their rationale, according to Tosteson, was that he had already made major contributions to the project and it was close to completion.

Late in September 1981, the Harvard team sent its portion of the AMPIM data to the NIH in preparation for a group meeting in Bethesda at which the participants would, for the first time, discuss unblinded results. (Darsee and Kloner in April had presented blinded results at an NIH meeting.) On 6 October, however, NIH called Harvard to say there were peculiar problems with the data from Braunwald's lab, and that the investigators should be prepared to discuss them at the group meeting on 30 October. Harvard at that time did not comment on the situation. It was not until 3 weeks later, just 2 days before the meeting, that Harvard confided to NIH officials that the trouble centered on an experimenter named Darsee. "They called to say they had a serious problem," says Jerome G. Green, head of extramural affairs at the heart institute, "and they were told, well, you're just going to have to go ahead and describe it and let it all hang out." At the meeting in Bethesda, lab chief Kloner laid out the whole story. Says Keith A. Reimer, a cardiologist from Duke who was at the meeting: "The data were too clean, and when you started plotting them to see what types of correlations would show up, based on the work of other labs, the correlations did not show. There were a number of inconsistencies. The Harvard people had no explanation, aside from the fact that John Darsee had been involved in the collection of the data."

The Harvard data also showed that the drugs under study—ibuprofen and verapamil—limited the damage to heart muscle when a coronary artery was blocked. Darsee, in fact, had published similar conclusions about ibuprofen in many of his abstracts and in the October paper he coauthored with Kloner. But while Darsee's data presented at the AMPIM meeting showed a protective effect for the drug, the data from other centers in the NIH study did not. The final problem, according to Green at NIH, was the basis for the Harvard conclusions. "They said at the meeting that part of the problem was the lack of raw data. They did not have the raw data for the blood flow." Such data were critical for the study, because blood flow showed whether the drugs allowed more blood to get to the heart muscle. Harvard withdrew the data from the study.

The news about Darsee was now starting to circulate in the biomedical community at large.

In November, 6 months after Darsee admitted the falsification, Harvard took the steps that might be expected to follow instantly upon the discovery of scientific fraud and the hint of widespread problems. It was too late to withdraw the six abstracts printed in the October issue of *Circulation*, but Darsee was prevented from presenting the full explanation of the abstracts at the AHA annual meet-

"After 18 months or so," says Braunwald, "it is not the custom to hold onto a fellow's hand at every turn."

ing, held from 16 to 19 November in Dallas. (He did, however, present papers at the meeting based on the four Emory abstracts.) And despite the new questions the AMPIM study had raised about the veracity of his work, he was still allowed to visit Braunwald's lab. In early December, Tosteson formed the blue-ribbon panel to look into the whole affair.

The attitude of Harvard officials during the unfolding of the affair had been based on the presumption of innocence. Darsee, a fellow with impeccable credentials and a bright future, had admitted to one act of fakery but denied any other improprieties. (Darsee, in a brief telephone interview, would neither confirm or deny this or any other point of the story.) The investigation, according to sources close to the lab, at first substantiated this faith. The first work they examined—that which had been carried out when Darsee first arrived in the lab and already had been published—looked perfect, raw data and all. It was only when Kloner and Braunwald began to go over more recent work that problems emerged with the raw data. "They began to find holes," according to one source, "and Darsee's defense was, well, you never told me to keep the data. And so they were left with a situation where it looked bad, but there was no definitive proof of systematic cheating." The period in which the volume of raw data starts to fall off is after Darsee had been at the lab for a little over a year—in other words, the point at which supervision relaxed.

Given the one admission of data fabrication, an observer might see the absence of supporting data for other research as shifting the burden of proof to Darsee. That is not what happened at Harvard. Tosteson attributes the delay to Braunwald's desire to "manage" the situation. "The record will show that Braunwald was extremely anxious to try and manage the situation in such a way as to fulfill the promise that he saw in Darsee," says Tosteson. "It has to do with the changing perceptions of the extent and nature of the falsification. It had to do with increasingly wide looks at what Darsee was doing." The question of innocence was rendered moot, however, when the AMPIM study compelled Harvard authorities to explain to NIH that they had a problem. Asked why the absence of raw data for the AMPIM study was not discovered earlier, sources close to the lab say that at that time of the NIH query—some 5 months after Darsee acknowledged the falsification—Kloner and Braunwald still had

not gotten around to thoroughly examining Darsee's work in progress.

Braunwald says he has no doubts about anything that has been published. "The *PNAS* paper in November or any other paper that bears his name, I have total confidence in," he told *Science*. The nagging question, in light of the slow pace of the investigation done in areas such as the AMPIM study, is how much weight to put on these assertions. Perhaps only a detailed audit will be able to resolve the issue.

Had the situation been addressed with vigor right at the start, perhaps such outside advice might not have been necessary. "These things have to be dealt with in a forthright way," says David Dressler, a member of the Harvard biochemistry department who in the early 1970's coauthored several papers with a junior researcher who later admitted cheating. The papers were promptly withdrawn. "To have this thing circle back from the NIH raises questions of institutional behavior. In my own case, I never thought about handling it through the university but informed the scientific community right away. . . . The idea of somebody coming down to see the original data, implying that papers had been sent off based on work the senior researcher had never looked at, is quite amazing to me."

A major mandate of the blue-ribbon committee* is to pass judgment on Harvard's handling of the affair. But the chairman of the committee, Richard S. Ross, dean of Johns Hopkins School of Medicine, has already expressed some views on the issue. Said Ross, after the committee held its first meeting: "The Harvard people have been anxious to do it the way they've done it, to avoid letting it smoulder like it did at New Haven. I mean they've come clean, and they've got some external people to look at it. The idea is that we are going to do it vigorously and let the chips fall where they may."

Late in December, NIH decided it too will investigate the Darsee affair. A panel of five or six outside experts will travel to Boston on behalf of NIH during the early part of 1982, according to Green. "It is virtually certain that we will tell Harvard to go ahead and repeat the [AMPIM] study at no cost to the government, but not until the investigations are over."

The Darsee affair is but one of several such incidents that have emerged with apparently increasing frequency during the past few years. One who has addressed the question of why such a trend may be occurring is Robert H. Ebert, former dean of Harvard Medical School and currently president of the New York-based Milbank Memorial Fund. "I don't condone this behavior," he says, "but certain things in our culture disturb me about all this. One is the enormous importance that is attached to the number of publications by committees that consider people for promotion. . . . And I don't want to be critical of any individual in this current affair, but I think that with the pressure to produce, the pressure for publication, I'm not sure the supervision is as close as it should be." Ebert also believes that such incidents are far more frequent than is often acknowledged. "There is a borderline falsification that is more common than anybody knows, in which you are anticipating the results you are going to get when you put in an abstract. That whole environment is bad. There should be such a great value put on accuracy that it would never occur to anybody to do that. It is kind of a moral issue of our times."

Perhaps Ebert is correct, and supervision is not what it has been in the past. The fellows in the lab at the Brigham were suspicious of Darsee's prodigious output, having sat next to him day after day. Their contentions of widespread fakery, however, at first met with a fair amount of skepticism. Perhaps, if senior investigators Kloner and Braunwald had been closer to the actual happenings in the lab, they too would have been more immediately concerned. Clearly, it was Braunwald's hope that Darsee would succeed at Harvard in a spectacular way. But such hopes are best fulfilled in an atmosphere of oversight and intellectual sharing that lies at the heart of the relationship between master and apprentice. As has been demonstrated in cases of dishonesty at other universities, a lack of guidance can cause such hopes to come crashing down in ruin. The various groups investigating the Darsee affair have a challenging mandate. They need to find not only why the incident at Harvard occurred, but how such episodes might be prevented in the future.—WILLIAM J. BROAD

A second story will report the conclusions of the blue ribbon committee investigating the Harvard affair.

*Also on the committee are A. Clifford Barger, professor of physiology at Harvard; Baruj Benacerraf, a professor of comparative pathology at Harvard and president of the Sidney Farber Cancer Institute; Burton S. Drebin, professor of philosophy at Harvard; Saul J. Farber, dean for academic affairs and professor and chairman, department of medicine, New York University School of Medicine; Gerald Frug, professor of law at Harvard; Robert I. Levy, dean of Tufts University School of Medicine; and Joseph B. Martin, professor of neurology at Harvard.

0036-8075/82/0129-0482$01.00/0 Copyright © 1982 AAAS

Report Absolves Harvard in Case of Fakery

But NIH official "disappointed" with panel's report chides Harvard on national television for delay in reporting incident of fraud

A blue-ribbon committee has absolved officials at the Harvard Medical School of any blame in allowing John R. Darsee to perform research for 6 months after he confessed last May to the fabrication of data. The eight-person committee,* appointed by Harvard Medical School dean Daniel C. Tosteson in December 1981 and chaired by Richard S. Ross of the Johns Hopkins School of Medicine, included five members of the Harvard faculty.

After the admission of fakery, Darsee was stripped of his appointment by his superior, Eugene Braunwald, but continued publishing papers and working in one of Braunwald's cardiovascular labs at the Brigham and Women's Hospital, on a $724,154 multicenter study financed by the National Institutes of Health (NIH). The lab at Brigham is run for Braunwald by Robert A. Kloner.

"The responses of the institution seem reasonable in relation to the information available at the respective times," says the committee in its 12-page report. "In retrospect, it is clear that Drs. Braunwald and Kloner felt in May and June that they were probably dealing with a single bizarre act by a young man who had performed exceptionally well previously. In light of this consideration, the plan selected in June seems to be reasonable in that it provided an opportunity to assess the extent of the damage and also to provide for a period of observation under supervision."

The committee also found reasonable Harvard's decision not to notify NIH in May that an admitted data fabricator was working on a federal study, but an NIH official did not. "I was disappointed," William Raub, head of extramural affairs for NIH, told *Science*, "that the report was not explicit about informing the sponsor." Although the Ross report does address the general question of who should have been informed, it makes no mention of federal collaborators. Raub appeared on a CBS news program to say that Harvard had an "obligation" to immediately inform NIH of the admitted fabrication. The committee did reprove Harvard's failure to notify Darsee's Harvard collaborators, but this tepid criticism was for the most part lost in the report's overall exoneration.

The 6-month "period of observation" revealed that two studies, including the NIH multicenter project, contained what the Ross report calls "unusual results which are highly suspect." Darsee in meetings with the committee denied any improprieties except for the one admitted fabrication in May.

According to the committee report, Braunwald decided that Darsee should continue to work on the NIH study "under close supervision" during the summer and fall for two reasons. "First, the study was heavily dependent on constant procedures and techniques, and a change in investigator in one of the collaborating laboratories might jeopardize the entire study. Second, the study was randomized, blind, and carried out simultaneously at four independent laboratories, and the results were analyzed centrally at the NIH. Therefore, any mishandling of data should be readily apparent from a comparison of results from different laboratories."

"On the other hand," notes the Ross report, "this course of action was not without cost. Darsee's presence in the laboratory environment after the events of May proved to be damaging to the morale and productivity of the laboratory."

Although as of 30 June Darsee was no longer officially a member of the Harvard staff, he was paid privately, according to the Ross report, from a "gift account" which Braunwald managed.

The silence of the Ross report on the issue of federal notification is curious because one area where the committee suggested "ways in which the institutional response could have been improved" was in the notification of Darsee's collaborators—but only collaborators at Harvard. "In May a systematic search should have been conducted to notify all persons within the institution with whom Dr. Darsee had collaborated," it said. The reason for the committee's mild reprimand is that at least one collaborator of Darsee's did not hear about the problems until November—some 6 months after Darsee confessed to the original fakery. In this collaboration—carried out with Leonard Holman of the nuclear medicine division at Brigham, a division Braunwald does not head—the committee found that the data "appear to have been manipulated." At a Harvard news conference held when the report was released on 25 January, Braunwald said he was not aware of Darsee's collaboration with Holman and therefore did not inform him of the problems.

Aside from the issue of collaboration, the blue-ribbon panel's only other point that bordered on criticism of Harvard's handling of the Darsee affair was that "a small committee of senior professors from within the University, but outside the involved department, should have been consulted immediately after the discovery in May. . . . Such a committee, be it standing or ad hoc, could have shared the burden with the Dean, the Chairman of the Department, and Laboratory Director and offered objective advice concerning management of the problem."

Not a hint of criticism is found in the report over the issue of supervision or laboratory standards. Two members of the blue-ribbon panel, Robert I. Levy and A. Clifford Barger, visited the Brigham laboratory on the morning before Christmas and subsequently reported to the full committee that "the present problem does not appear to be at all referable to the existing Cardiac Research Laboratory standards, policies or procedures nor to overt pressure provided by its Director or Dr. Braunwald. Dr. Robert Kloner and the Cardiac Research Laboratory have maintained an extremely effective system for data collection, analysis and storage."

On a related issue, the report noted

*Richard S. Ross, chairman and dean of Johns Hopkins School of Medicine; A. Clifford Barger, professor of physiology at Harvard; Baruj Benacerraf, professor of comparative pathology at Harvard and president of the Sidney Farber Cancer Institute; Burton S. Dreben, professor of philosophy at Harvard; Saul J. Farber, dean for academic affairs at New York University School of Medicine; Gerald Frug, professor of law at Harvard; Robert I. Levy, dean of Tufts University School of Medicine; and Joseph B. Martin, professor of neurology at Harvard.

0036-8075/82/0212-0874$01.00/0 Copyright © 1982 AAAS

that raw data were missing for each of the suspect studies. However, it does not speak directly to the overall issue of whether the original data for many of Darsee's dozens of other experiments at Harvard were unaccountably missing—a contention of sources quite close to the laboratory (*Science*, 29 January, p. 478).

Nonetheless, the nearly 100 papers and abstracts published by Darsee were judged by the committee to be in order. "A systematic review of all work involving Darsee has been conducted by Drs. Kloner and Braunwald. Primary data have been reviewed and Darsee's co-workers have been questioned. The committee is convinced that Drs. Braunwald and Kloner have documented the extent of the irregularities and properly concluded that the previously published work from the laboratory in which Darsee was a member of the group is accurate. None of the work under suspicion has been published. The papers and abstracts containing fabricated data have been withdrawn."

The report makes no mention of the possibility of future publications. A review of the literature, however, might lead one to believe that Darsee will continue to present papers based on his work at the Brigham. The advance program for the 31st annual Scientific Session of the American College of Cardiology lists an "original contribution" from Darsee and Kloner, scheduled to be presented at the Atlanta meeting in April 1982. Braunwald, however, told *Science* that this and any other work by Darsee that was in the mill has been withdrawn.

Another issue not addressed by the committee in its report is the fact that the colleagues of Darsee's who witnessed the "single bizarre act" soon afterward told Kloner and Braunwald that they suspected Darsee was systematically faking a major part of his prodigious output. According to chairman Ross, the committee did not meet with any of these colleagues.

At the Harvard news conference, Braunwald emphasized the lack of evidence of widespread problems in May and explained why he chose not to make the case public. "At the moment we had a brilliant person," he said. "He clearly was one of the most outstanding, or the most outstanding, of the 130 research fellows I have been privileged to work with. Public disclosure would have ruined him for life. We felt that to do this on the basis of a single incident would have been extreme. So we took a guarded position."

"Institutional Response"

The majority of the section of the Ross committee report that deals with Harvard's "institutional response" follows:

The committee examined the response of the Harvard Medical School to the Darsee incident from the date of discovery on May 22 to the present time. The purpose of this examination was to determine whether any future steps should be taken, whether Dr. Darsee had been treated fairly and what lessons might be learned with regard to the handling of such unfortunate events if they should occur in the future.

The response of the Harvard Medical School should be considered in two phases in accordance with the information available at that time. The first phase response began in late May and early June, following discovery and admission of data falsification by Dr. Darsee. Dr. Darsee's fellowship was terminated on June 30, and his faculty appointment, which had been offered for the following year, was withdrawn. The Dean's Office was notified, and abstracts of work in question were not submitted. Unfortunately, Dr. Holman, Dr. Darsee's collaborator in the Division of Nuclear Medicine did not learn of the fabrication of data in the "Extension vs. Expansion Study" until November and was not informed that Dr. Darsee's appointments had been terminated.

The second phase response occurred in October and November and was initiated by the discovery of the irregularities in the data [for two studies]. At this point in October and November, when it seemed likely that more than one episode was involved, the granting agency [the National Heart, Lung, and Blood Institute] was notified and two manuscripts . . . were withdrawn. Dr. Darsee agreed not to present any papers originating from the Harvard laboratories at the annual meeting of the American Heart Association. Officials at Emory University were notified that there were questions about Dr. Darsee's work. . . . In November, the Dean of the Harvard Medical School, Dr. Tosteson, decided that the matter was of such importance that a committee from within and outside Harvard University should be appointed. . . .

The committee questioned the reason for the delay between the initial event and response in May and June and the full-scale investigation in November. In retrospect, it is clear that Drs. Braunwald and Kloner felt in May and June that they were probably dealing with a single bizarre act by a young man who had performed exceptionally well previously. In the light of this consideration, the plan selected in June seems to be reasonable.

A systematic review of all work involving Darsee has been conducted by Drs. Kloner and Braunwald. Primary data have been reviewed and Darsee's co-workers have been questioned. The committee is convinced that Drs. Braunwald and Kloner have documented the extent of the irregularities and properly concluded that the previously published work from the laboratory in which Darsee was a member of the group is accurate. None of the work under suspicion has been published. The papers and abstracts containing fabricated data have been withdrawn.

The committee believes Dr. Darsee has been treated fairly. For several months he was dealt with in a manner consistent with his assertion that he was guilty of only one isolated act of data fabrication. The committee approves of the conscientious and responsible attitude of Drs. Braunwald and Kloner in this troublesome affair.

The committee suggests that there are two ways in which the institutional response could have been improved: First, a small committee of senior professors from within the University, but outside the involved department, should have been consulted immediately after the discovery in May. In any case, such a committee, be it standing or ad hoc, could have shared the burden with the Dean, the Chairman of the Department, and Laboratory Director and offered objective advice concerning the management of the problem. The second suggestion has to do with internal communications. In May a systematic search should have been conducted to identify all persons within the institution with whom Dr. Darsee had collaborated, and these persons should have been informed confidentially. . . .

NIH officials stress that their only criticism of the report and Harvard's behavior has to do with the issue of federal notification. Any other possible comments will wait, they say, until the NIH investigation of the affair is finished. A four-member team† of advisers to NIH will meet in February for initial review of the case and then travel to Harvard in March to conduct an investigation. NIH has not yet decided whether the report will be made public.

†Howard E. Morgan, chairman of the committee and professor of physiology at Pennsylvania State University; Francis J. Klocke, professor of medicine at the State University of New York at Buffalo; John T. Shepard, dean of the Mayo Medical School; C. Kern Wildenthal, dean of Southwestern Medical School, University of Texas at Dallas.

On the issue of notification, however, the NIH has taken a strong stance. Raub appeared in a live interview on the news program "CBS Morning with Charles Kuralt and Diane Sawyer" the day after Harvard released the blue-ribbon committee's report. Asked if Harvard or any medical center using federal funds had "a legal or at least an ethical or moral responsibility to notify NIH when something like this happened," Raub replied: "They do have an obligation, and we in turn have an obligation to the public whose tax money this is to ensure that every case of real or apparent misdoing in science is identified and dealt with promptly and forcefully." Raub later told *Science* that in the Harvard case the prompt notification of federal officials applied even if the contentions of Darsee's colleagues were ignored or the questions about the retention of raw data were left unanswered. "Where an incident is witnessed, where it is admitted—that is enough to go over my threshold," says Raub, who spoke on the subject for NIH at a 1981 congressional hearing on fraud in biomedical research. "The event itself is then worthy of notification. On the other hand, reporting it doesn't necessarily mean banishing the guy into the outer darkness. I think they could very well expect the feds to participate in a positive effort to try to help deal with the individual."

—WILLIAM J. BROAD

"Prevention of Dishonesty in Science"

In addition to evaluating Harvard's response to the recent incident of fraud, the blue-ribbon committee offered suggestions about preventing dishonesty in research and also called for a national conference to consider several important questions about the scientific environment and the responsibilities of institutions in which dishonesty occurs. The essence of their statement follows:

. . . The committee believes that the preventive measures most likely to be effective are those directed at the circumstances and practices in the laboratory which stimulate dishonest behavior in some people. In addition, observing the behavior of investigators working in the laboratory can be helpful.

The committee observed that a reported major scientific discovery will be confirmed or refuted rapidly because of the great impact it will have on the direction of science and the strong motivation of other laboratories to verify important data. The probability that spurious or fraudulent results will be discovered by others relates directly to the significance of the report and the number of scientists working in the same field. Therefore, the practice of reporting research results in small segments affords an opportunity for dishonesty because such reports are less likely to be verified by others.

The following measures are suggested as helpful in preventing dishonesty and its detection if it does occur:

- Special attention should be given at the time of recruiting to the motivation and integrity of the applicant through careful examination of credentials and claimed accomplishments.
- Written, detailed, explicit procedures for data gathering, storage, and analysis are essential and should be available in all laboratories.
- Fellows should be supervised by experienced scientists in the laboratory including regular, in-depth scrutiny of the primary data and the calculations leading to the presentation of results. The laboratory director should, by example, develop in the fellows a respect for primary data and its preservation.
- The conduct of studies which are blind or coded, and the exchange of reagents or methods between laboratories should be encouraged. The repetition of research in the same laboratory and by different research teams is desirable whenever possible.
- Fellows should be encouraged to work with other colleagues, to share data and to engage in free discussion of results. Secrecy about methods and data should be discouraged.
- Emphasis should be placed upon the quality and significance of research rather than on quantity and visibility. The laboratory director should accept responsibility for the quality of the work reported from the laboratory and should actively discourage the submission of multiple abstracts to national meetings. Abstracts should be considered as reports of results worthy of publication and as forerunners of complete papers rather than as bids for appearance on a national program.
- There should be close personal interaction between faculty and fellows. Among the many benefits of such close interaction could be the early detection of personal problems or unusual personality traits.

We further suggest that a national conference involving the NIH and universities be convened to consider a number of unanswered questions:

- Are there steps which could be taken at the national level to encourage the evaluation of young scientists on the basis of quality rather than numbers of publications? These policies must be addressed not only by universities, but also by editorial boards, granting agencies, and study sections. It has been suggested that scientists who fragment their results into many papers should be criticized rather than rewarded.
- Should national societies examine their policy regarding limitation of multiple abstracts with the same first author?
- What is the responsibility of the institution discovering dishonesty in scientific research to other institutions, to the scientific and medical community, and to the public at large?

Coping with Fraud: The Darsee Case

"New evidence suggests that papers published in refereed journals contain fabricated data from Emory as well as Harvard"

It has been nearly 2 years since John R. Darsee was caught fabricating data in an experiment at the Harvard Medical School, where he was a fellow in the Cardiac Research Laboratory headed by Eugene Braunwald. The Darsee case is one of a handful of well-publicized incidents of fraud that have occurred within the past few years at some of the country's most prominent academic institutions. In their wake, unprecedented attention is being focused on the causes of research fraud and the way institutions respond to evidence that someone has fabricated results.

Darsee was caught in the act of falsifying data on 21 May 1981 during an experiment on dogs in Braunwald's laboratory at the Brigham and Women's Hospital, a Harvard affiliate. The incident precipitated a complex and costly series of responses, first at Harvard and subsequently elsewhere. After privately notifying Harvard officials, Braunwald and his principal research deputy, Robert Kloner, began an internal review of Darsee's research that would go on for more than 1½ years. In late November 1981, after it appeared that Darsee's fraud extended to a major collaborative study sponsored by the National Institutes of Health (NIH)—on which Darsee continued to work even after he was caught fabricating data—the dean of Harvard, Daniel C. Tosteson, asked the dean of Johns Hopkins, Richard S. Ross, to head a committee to investigate the affair. In December, the director of the National Heart, Blood, and Lung Institute which was funding most of Darsee's research, appointed four university-based cardiologists to conduct an independent audit for the government. That panel, chaired by Howard E. Morgan of the Pennsylvania State University College of Medicine at Hershey, spent 6 months on the case.* Then, NIH director James B. Wyngaarden instructed one of his deputies, William Raub, to convene a group of senior NIH staff to review the Morgan panel's report.

Wyngaarden accepted their recommendation that Darsee be debarred from receiving NIH funds or sitting on advisory bodies for 10 years (*Science*, 25 Feb., p. 937). Formal debarment proceedings are now under way in what may be the harshest sanction NIH has ever levied against a scientist. The case has been referred to the Justice Department.

Meanwhile, back at Harvard, medical school faculty, working with the university's general counsel, were drafting guidelines about what to do if a Darsee-like incident were to happen again. A policy statement was released in February when the Morgan and Raub reports became public.

Eugene Braunwald
"I got a bum rap."

One might reasonably conclude from all this that the matter is finally settled. Not so. Still to be resolved is whether, and to what extent, Darsee committed fraud during his years of training at Emory University in Atlanta and as an undergraduate at Notre Dame. The answer to that affects the validity of several published papers coauthored with senior members of Emory's faculty, and it also has a bearing on a question that has dominated the case, namely, whether the pressures to produce that exist in Braunwald's lab were a major element in the data fabrication. It also raises difficult questions about Darsee's fitness to practice medicine. Darsee is currently a fellow in critical care medicine at Ellis Hospital in Schenectady, New York, where administrators are aware of his Harvard record and are standing by him.

Darsee's first known act of fraud—the 21 May incident—involved the labeling of data that had been obtained on dogs over a period of a few hours to make it look as if the data had been recorded over 2 weeks. It came to light only because he virtually allowed himself to be caught by fabricating tracings in plain sight of co-workers (*Science*, 29 Jan. 1982, p. 478). When confronted by Braunwald and Kloner, Darsee admitted the fabrication and said he was sorry, all the while insisting that his was a single, foolish act of misconduct. Braunwald believed him.

Of some 130 fellows who had passed through his laboratory over the years, Braunwald judged Darsee the most remarkable. He was a tremendously hard worker, but, says Braunwald, "He was not a drone." Rather, he thought of him as "brilliant" and "creative," a young man with whom he could discuss research projects as a scientific equal. Darsee was well-liked at Harvard, as he had been at Emory. In fact, on the very days Darsee was facing charges of fraud, his former teachers at Emory were writing Braunwald glowing letters of recommendation in support of Darsee's pending appointment to the Harvard Medical School faculty. Cardiologist Nanette K. Wegner, who coauthored papers with Darsee, wrote that, in addition to having superb research skills, "he enjoys an

*Other panel members were cardiologists Francis J. Klocke, State University of New York at Buffalo; John T. Shepherd, Mayo Medical School; and C. Kern Wildenthal, Southwestern Medical School, University of Texas Health Science Center at Dallas.

31-35

enviable relationship with his patients and with his peers . . . and in general is an extremely well-rounded human being." Another coauthor, associate dean Donald O. Nutter, spoke of Darsee's "intelligence" and "moral character."

By the time these letters were in Braunwald's hands, Darsee's offer of a faculty position at Harvard had been withdrawn and his NIH fellowship was being terminated. But Braunwald decided to keep him on in the laboratory to complete his part in the NIH collaborative study. This decision was subsequently criticized by NIH officials and Morgan panel members, who believe that Darsee's collaborators at other institutions had a right to know what was going on. But at the time, Braunwald was simply not prepared to believe that his star fellow had done anything more than crack—once—under pressure. As a friend of Braunwald's told *Science*, "Braunwald thought that maybe he'd ruined Darsee."

Harvard's initial institutional response was to handle the matter in confidence. No effort was made to identify and notify all of Darsee's co-workers at Harvard. NIH was not notified of the one admitted act of fraud. Letters from Darsee and the Brigham and Women's Hospital to NIH, terminating Darsee's fellowship did not tell NIH why.

Braunwald and Kloner spent months reviewing Darsee's notebooks but found nothing. For a while, it looked as if he was, in fact, guilty of only a single act. But in the fall, when NIH broke the code on data from the multi-institutional study Darsee was part of, it became clear the damage might be more widespread. Data from Braunwald's lab were inexplicably at variance with those from the collaborating groups. It was at this point that Tosteson asked Ross of Hopkins to head a committee of five Harvard faculty and three outsiders to review both Darsee's work and Harvard's handling of the matter. The Ross committee was appointed the day before Thanksgiving and reported 2 months later (*Science*, 12 Feb. 1982, p. 874). It was generally sympathetic to Braunwald and Tosteson, taking the view that they acted reasonably, given their information at the time. The committee, without conducting a serious investigation of its own, took the position that "Drs. Braunwald and Kloner have documented the extent of the irregularities in Darsee's data. . . ." In fact, it would be months before the full extent of the fraud became clear. In retrospect, Ross realizes it was unwise to succumb to Tosteson's urging that his committee conclude its business quickly. In the end, the Ross report contributed little to a resolution of the matter but did make clear the mistake of trying to conduct a serious investigation on the run.

The Morgan panel did conduct a rigorous statistical analysis of data in Darsee's Harvard publications and uncovered "extensive irregularities" in five papers on which he was first author, with Braunwald and Kloner as coauthors. "The primary responsibility for this situation rests with Dr. Darsee, but must be shared by his coauthors," the panel said. Braunwald and Kloner retracted findings

Howard E. Morgan

Headed inquiry for NIH that criticized most of those involved with the Darsee case.

in those papers that were based on Darsee's work.

It was while the Morgan panel inquiry was going on that hard, physical evidence of Darsee's data fabrication in the NIH study was discovered—ironically enough by Braunwald and Kloner. Analysis of heart tissue from experimental dogs which should have been injected with radioactive microspheres revealed no residual radioactivity in animals studied prior to 21 May. However, there was radioactivity in heart tissue from dog experiments Darsee conducted during the time he was in the lab after the 21 May incident. The "smoking gun" was thus in hand.

The Morgan panel's review of the Darsee case took it beyond the specific facts of data fabrication into the more tenuous realm of an assessment of the way Braunwald runs his laboratory. Where there is general agreement that Harvard's institutional response was flawed, there is disagreement over the question of Braunwald's supervision of his young researcher and the pressures in his lab to produce. The Morgan panel does not go so far as to blame Braunwald for Darsee's fraud, but it does fault his supervisory style for letting it go undetected. According to its report, "The Panel is of the opinion that the circumstances prevailing throughout Dr. Darsee's period in the laboratory, while not responsible for or in any way condoning his misdeeds, helped to create an environment that may have inhibited their being uncovered."

Says Braunwald, "I got a bum rap."

Eugene Braunwald is a renowned and powerful cardiologist who is widely credited with major discoveries in his field. In addition to directing a laboratory where he has four research fellows, Braunwald is in the unusual position of being chairman of medicine and physician-in-chief at two Harvard hospitals—the Brigham and Women's and the Beth Israel. He is not in the lab every day. Although he met with Kloner twice weekly, his meetings with the fellows were often scheduled 6 to 8 weeks or more apart.

According to Morgan, the panel felt that "there just wasn't enough direct contact with Braunwald." Their opinion, he says, is based not only on the Darsee case but also on interviews with fellows in the Harvard lab and on panelists' own supervisory practices. Morgan, for instance, says, "I see my people every day and have formal presentations of work every Monday. But then, I'm not running two departments of medicine."

Braunwald, who feels deeply stung and angry about these criticisms, believes that he is being unfairly scrutinized in that any number of other labs are run like his. In an interview with *Science* and in letters to NIH, Braunwald observed that it has been his practice for years to delegate authority in the research laboratory to a bright younger scientist and that it has never been an issue before. "One of the key roles of a Principal Investigator is to provide scientific leadership, ideas and, when possible, even inspiration," Braunwald said in a long memorandum to NIH officials. "Since 1960 I have maintained my direct participation in the laboratory by trying to offer these while at the same time providing an extremely capable colleague, such as Dr. Kloner, for day-to-day, on site, full-time participation in the actual experiments."

Although Raub's review committee of senior NIH staff took clear positions on most of the comments and recommendations of the Morgan panel, it ducked on this question of the way a laboratory should be supervised. In his official memorandum to NIH director Wyngaarden, Raub said, "The staff committee recognized that the Panel's findings in

this regard are matters of professional judgment and that equally qualified observers might form different impressions from the same information." The decision, therefore, was to make the panel's observations and the rebuttals from Braunwald and Kloner part of the public record.

One of the principal defenses Braunwald makes of his laboratory rests on his belief that Darsee has a long record of scientific deception which he managed to conceal from his superiors at Emory well before he arrived at Harvard. In March 1982, while the Morgan panel was conducting its investigation of Harvard, Braunwald started to review several of Darsee's papers from Emory. Almost immediately, he developed "concerns" about the integrity of three of them. He reported this to NIH and to Neil Moran at Emory, who had been named chairman of an internal review of Darsee's work there. Subsequently, Braunwald has detected possibly fraudulent data in more than a dozen of Darsee's papers from Emory. Several involve studies of human subjects, including patients in coronary distress.

It is Braunwald's contention that the Morgan panel should have expanded its investigation to include possible fraud at Emory before making judgments about the laboratory environment at Harvard. However, the panel declined to do so and made no mention of the allegations.

The decision to ignore that issue was made jointly by the Morgan panel and NIH staff for a number of reasons. First, technically, the Morgan panel was sponsored by the heart institute, and no heart institute funds were used to support Darsee at Emory. Second, Morgan notes that the panel was not prepared to make the necessary commitment of time to "become full-time investigators." Third, as Raub told *Science*, there was a feeling that the Morgan committee should "wrap things up" on the Harvard inquiry. Furthermore, as he said in his report to Braunwald, "We . . . believe . that there is ample basis to initiate debarment proceedings [against Darsee] now."

Nonetheless, NIH officials are now persuaded that the situation at Emory is sufficiently serious to warrant an NIH investigation "as soon as resources permit." Meanwhile, Raub says, NIH is standing by while the Moran committee completes its internal inquiry at Emory.

The Moran report is expected "within four weeks," according to Michael Aycock, assistant to the dean at Emory. When complete, the report will be reviewed by a panel of outside scientists appointed by Emory and by NIH.

Harvard Acknowledges It Could Have Done Better

Harvard has been severely criticized for its handling of the Darsee case by the Ross committee, the press, NIH officials, and, most recently, the Morgan panel. Each acknowledges that Harvard officials acted from honorable motives, but each concludes that the decision not to notify NIH promptly of Darsee's admitted act of fraud on 21 May 1981 and not to call immediately for an audit by persons outside Braunwald's lab was a mistake.

In a memo to the Harvard Medical School faculty timed to coincide with release of the NIH reports, Tosteson acknowledged his error. "A committee of persons not associated with the Cardiac Research Laboratory should have been appointed to investigate the matter and audit Dr. Darsee's research when his misconduct was first discovered in May, 1981," he said. "Also, a prompt and thorough effort should have been made to identify and notify other investigators. . . . I wish, in retrospect, those actions had been taken."

Braunwald, too, would do it differently another time. "I no longer think you can give someone the benefit of the doubt in any case where fraud has been admitted," Braunwald said in a lengthy interview with *Science*. "The cost to the institution and to science is too high. I have changed my opinion on this." One reason Braunwald chose not to notify NIH 2 years ago was fear that a government agency, bound by the Freedom of Information Act, could not keep the case private. Now, he says, "We can't afford confidentiality as far as NIH goes." Rather, "You have to have some way of issuing a clarion call of innocence if someone is the victim of crank allegations."

Harvard's new policy on "principles and procedures for dealing with allegations of faculty misconduct" constitutes recognition of the fact that its response to the Darsee case leaves something to be desired. And yet it would be wrong to say that its take-home lesson is that these things can be handled only one way. The policy, which vests authority for fraud investigations in the office of the dean, is designed to "permit flexibility" by giving him maximum discretion to "attempt to resolve the matter through informal processes and discussions" at one end of the spectrum to full notification of NIH and other interested parties at the other. It also leaves to the dean's judgment the decision on appointing a committee from outside the affected department. This stands in contrast to the Morgan panel's recommendation that such a committee be named "immediately."

The Association of American Medical Colleges made a similar recommendation last summer when it issued general guidelines on coping with fraud (*Science*, 16 July 1982, p. 226). Just how many research institutions have adopted guidelines of any sort is not known.

Although there are guidelines to advise institutions on procedures to follow when fraud is alleged, there are as yet none that spell out just how the investigation itself should be conducted. Braunwald's attorney, Bancroft Littlefield, thinks that "doctors aren't equipped by training" to conduct these kinds of investigations themselves. Although innumerable people were reviewing Darsee's data, it was almost a full year after the inquiry began before Kloner found the "smoking gun" when he analyzed heart tissue from a dog in which Darsee had allegedly implanted radioactive microspheres. No radioactivity was present in the heart.

Says Littlefield of the investigations, "Obviously at first Braunwald and Kloner didn't do it well. The Ross committee didn't do well. And Morgan and company didn't find the smoking gun even with the help of NIH staff and access to all the material." The reviews tended to focus on the contents of laboratory notebooks, for instance, rather than an analysis of physical evidence. "What doctors need," says Littlefield, "is a manual." The first thing it should say, he argues, is that the laboratory should be isolated and the data locked up. NIH officials and others who have been involved in investigating allegations of fraud agree that there is a problem here—that review committees need what one of them has called "professional sleuthing help" in addition to scientific expertise.—B.J.C.

Although Emory is remaining silent, some indication of where the Moran committee stands is evident from correspondence released by NIH. Clearly, Moran is finding problems. According to a 13 December memorandum from Braunwald to Raub, in October Moran informed Braunwald that "his committee had discovered serious problems regarding Dr. Darsee's work at Emory." Raub also got similar word from Moran and told Braunwald that "the NIH staff's preliminary review confirmed the existence of such problems."

Recently, evidence has come to light that suggests the trail of Darsee's misconduct goes back to his undergraduate days at Notre Dame, where microbiologist Julian Pleasants heard him give a student seminar in 1969. Prompted by news accounts of the Darsee case, Pleasants looked up two papers Darsee published in the student-run *Notre Dame Science Quarterly*. One paper, on hormones and aging, describes an experiment in which blood was drawn from the tails of 200 rats weekly for the animals' lifetime of 90 weeks or more. In a letter to Braunwald, Pleasants said "By internal evidence these articles are fabrications." Pleasants told *Science* he decided to write because "I felt Braunwald was being blamed unfairly for having changed a person's character when it was already set. It was unfair to Dr. Braunwald and to the general practice of research."

Darsee's response to these various allegations is ambiguous. Although he clearly admits fabricating data in the 21 May incident, he seems to have denied responsibility for other misconduct. According to the Morgan panel report, Darsee responded to written questions about his data and also brought his attorney, Robert Gerard of Boston, to Bethesda for a 3½-hour interview. This, the report says, "provided Dr. Darsee with a full opportunity to present his own account of events and circumstances." In particular, the panel sought explanations for its discovery of statistical aberrations in tabular data from five papers on which he was first author. No primary data were available for review but Darsee did offer "suggested explanations for the surprising statistical characteristics of his data." The Morgan panel found his explanations "unconvincing."

Darsee also wrote a letter to NIH that Raub found so "highly personal" in its references to the death of his father and his admiration for Braunwald that he has acceded to Darsee's request that it not be released. However, in his memorandum to Wyngaarden, Raub wrote about "a letter from Dr. Darsee stating that, although he had no recollection of falsifying any research data, he acknowledged that the Panel's inquiry had established both the fact of falsification and his personal role. . . . It should be noted that Dr. Darsee's letter was a significant departure from his earlier assertions to the Panel that he had not engaged in any deceptive or irregular practices other than the single incident in May 1981."

Darsee is not speaking to the press, and attorney Gerard told *Science* that he cannot elaborate on Darsee's letter because "I don't know [what incidents] Darsee's statement about fraud referred to. He drafted the statement himself." Says Gerard, he does not "admit" to fraud "but says he doesn't recall."

Darsee has made one public comment on the case, which was released by a spokesman for Ellis Hospital after the Morgan and Raub reports were published in February. "I am asking forgiveness for whatever I have done wrong, and want to contribute to the medical system," Darsee said. "I must take responsibility for my actions and realize that it is my fault and no one else's."

According to hospital spokesman Andrew Foster, Darsee told hospital administrators about the situation at Harvard when he sought a fellowship in critical care medicine. His work there is "carefully supervised," Foster said, adding that Ellis Hospital believes his potential contributions to medicine and the Schenectady area are "great."

Harvard officials also notified Ellis Hospital about the Darsee case, by letter when they first learned he was working there and a second time by telephone prior to the release of the Morgan and Raub reports. NIH has made no formal notification to Ellis, Raub says, but is aware that Harvard officials notified the hospital of the debarment proceedings. "Knowing what he knows now, hospital administrator William Schirmer would hire him again," Foster told *Science*.

Until the 21 May incident at Harvard, Darsee was considered one of the brightest young cardiovascular researchers in the country. His papers, coauthored by major figures in the field, were published in important, refereed journals. And now, as those papers are reread with what Braunwald calls the "Sherlock Holmes perspective," it seems that they are full of errors. What is now apparent is that the data, in general, are too good, too neat, too perfect to be believed.

Why did no one catch it before? Braunwald, for his part, admits that even in the areas he knows best he failed to see what now seems obvious. "It takes a different mind set," he says, "to look at a paper and think total fraud." Says Morgan, "Maybe cases such as this will increase referees awareness when they're reviewing a paper." Maybe so.

—BARBARA J. CULLITON

Medical News

April 8, 1983

The Harvard fraud case: where does the problem lie?

A scientific scandal at Harvard Medical School has culminated in the harshest penalty the federal government has ever imposed on a researcher for falsifying data.

Moreover, a National Institutes of Health (NIH) investigative report issued Feb 16 and accompanying documents have criticized Harvard for its handling of the affair, embarrassed one of the nation's leading cardiologists, and suggested strongly that the pattern of fraud extends back to the perpetrator's residency days at Emory University School of Medicine in Atlanta.

The protagonist in the case, one of the most complex in the expanding annals of scientific fraud, is a former Harvard research fellow, John Roland Darsee, MD, 34. Now in a second residency program in upstate New York, Darsee appeared destined for a brilliant academic career before his boldness in faking results in animal experiments began to reveal an extensive pattern of data fabrication. Ultimately, he was found to have falsified large amounts of data in at least five Harvard research projects with which he was involved, all animal studies.

The bogus data found their way into nine coauthored published papers as well as numerous abstracts, published and unpublished, and they contributed greatly (along with other, unrelated factors) to the ruin of a $724,154, multicenter, cooperative study done under contract with the National Heart, Lung and Blood Institute (NHLBI) to devise animal models for assessment of therapies designed to protect ischemic myocardium (the AMPIM study). A wide range of Darsee's work at Emory, where he excelled during five years of postgraduate training, is also under suspicion, and the behavior pattern may even extend to Darsee's undergraduate college days.

Following a five-month NHLBI inquiry into the problem by a panel of outside experts and another 7½ months of internal deliberations on the panel's report, the NIH has barred Darsee from receiving funds or participating in NIH committees or study sections for ten years. The NIH has also asked Harvard-affiliated Brigham and Women's Hospital to pay back $122,371, the funds it received for the AMPIM study. According to Mary L. Miers of the NIH's Office of Extramural Research and Training, this is the first time that the agency has demanded repayment of funds because a study was tainted by fraudulent data.

In addition, NIH Director James B. Wyngaarden, MD, has put the Cardiac Research Laboratory of Harvard Professor Eugene Braunwald, MD, where much of Darsee's cheating took place, on one year's probation to ensure that the laboratory maintains adequate standards of record-keeping and supervision before further federal funding is approved.

Braunwald apparently instigated the investigation at Emory. In an attempt to show that Darsee's fraudulent behavior was not linked solely to conditions at his Harvard laboratory—as the NHLBI panel suggested—Braunwald last year claimed that at least three of Darsee's Emory papers, which had been published in the *American Journal of Cardiology* and the *New England Journal of Medicine*, bore the earmarks of cheating, too. According to Braunwald, Emory's ongoing internal investigation, under the direction of Neil C. Moran, MD, professor of pharmacology, has thus far called into question at least six published papers and "a large number of abstracts."

Though the NIH decided that it already had enough evidence to proceed against Darsee, William F. Raub, PhD, NIH associate director for extramural research and training, wrote Wyngaarden a decision memorandum in January saying in part that "the consensus now is that the alleged misconduct at Emory is very serious and that NIH should undertake an investigation at Emory as soon as resources permit." Raub said in a recent interview that the NIH will await the results of Emory's inquiry before proceeding with its own. An Emory spokesman told *JAMA* MEDICAL NEWS that the school's internal investigation will be completed very soon but will be subject to external review before a report is made public.

Dr Braunwald

Darsee, who has left research and is retraining as a critical care specialist at Ellis Hospital in Schenectady, NY, has shunned all interviews since the scandal first became public in December 1981. He had confessed immediately to a single act of data fabrication (doing blood flow studies on a dog for several hours and then mislabeling the strip chart recording so that

continued on next page

continued from previous page
it would appear that the data covered a period of two weeks) in May 1981, after being confronted with the eyewitness testimony of two co-workers. But for more than a year afterward he insisted that it was an isolated, impulsive act, even as evidence of more systematic cheating began to accumulate. At one point last spring, according to Braunwald, he threatened to sue Braunwald but has not pursued that course. In March 1982, his Boston attorney, Robert C. Gerrard, said in a telephone interview that Darsee "has been savaged at all turns" and "has been had by events." Shortly after that, however, Darsee wrote a confidential letter to NIH officials in which he stated that he had no memory of falsifying research data apart from the one incident but acknowledged that the NHLBI's investigation established that he had. When the NIH announced its decision to bar him from further funding, he released a brief statement through the Schenectady hospital asking for "forgiveness for whatever I have done wrong."

Spreading the Blame

As noted, the NIH's action in the Darsee case constitutes the government's most severe response yet to an instance of scientific misconduct (as distinguished from fiscal malfeasance by grantees and contractors). Repeated revelations of research fraud in recent years (at Boston University Medical Center, Massachusetts General Hospital, Mount Sinai School of Medicine, Yale University School of Medicine, Sloan-Kettering Memorial Cancer Center, and elsewhere) have put the NIH under increasing congressional pressure to get tough with transgressors—or face enactment of legislation that will do the job. In response, federal health officials are in the final stages of formulating guidelines on scientific misconduct that will apply throughout federal health agencies.

Edward N. Brandt, MD, assistant secretary for health in the Department of Health and Human Services, has been speaking out on the subject as well. "Until recently," Brandt wrote last December in *Clinical Research* (1982;30:433-434), "we have been able to treat each incident as an isolated event. It is clear that more explicit and predictable procedures are needed if we are to exercise proper stewardship of public funds. . . ."

The NIH decision in the Darsee case also stands out because, for the first time, the agency has extended its conclusions and sanctions beyond the individual researcher found guilty to the scientists in charge of the laboratory where it took place. A panel of four outside experts appointed by the NIH to look into the affair* commented in its 36-page report on the Harvard laboratory's "hurried pace and emphasis on productivity, coupled with limited interaction with senior scientists." The report concluded: "The supervisory practices in the laboratory, while in no way responsible for Dr Darsee's misconduct, may have contributed inadvertently to the ease with which he was able to produce fabricated data and to the subsequent difficulty in documenting the extent of the problem."

Not only did the NIH-sponsored panel decide that accountability was not all it should have been at Harvard's Cardiac Research Laboratory, but it also rebuked both Braunwald and Harvard Medical School Dean Daniel C. Tosteson, MD, for their handling of the matter once the first documented instance of data fabrication by Darsee came to light.

"The report states, and we agree," Tosteson responded in a letter to the NIH, "that despite a good faith effort, the Brigham and Women's Hospital and the Harvard Medical School made mistakes in handling the case." Tosteson conceded that an independent panel of investigators should have been appointed promptly to determine the extent of the fraud, instead of permitting Braunwald and his deputy, Robert A. Kloner, MD, PhD, to handle it alone. And, the dean agreed, the NIH should have been informed immediately when Darsee was caught fabricating data. Instead, Braunwald and Kloner watched him more closely for five months and began studying his data. Then, as evidence mounted that fraud was extensive, Tosteson appointed a "blue ribbon" panel in November 1981.

However, rather than conducting its own scientific audit, the Harvard committee accepted the conclusion of Braunwald and Kloner that the fabrication was limited to the NHLBI contract study and three others, none of them published. In its report of January 1982, the blue ribbon committee pronounced itself "convinced that Drs Braunwald and Kloner have documented the extent of the irregularities and properly concluded that the previously published work from the laboratory in which Dr Darsee was a member . . . is accurate." Those words would later come back to haunt Harvard and its committee. The NHLBI investigators said that Harvard's investigation was a hurry-up job and that the Braunwald-Kloner audit it relied on "was insufficiently rigorous and not definitive."

Braunwald and Kloner, stung by the NHLBI panel's criticism, fought unsuccessfully throughout late 1982 to persuade the NIH to expunge from its final report any statements either implicating lax laboratory supervision as a factor in Darsee's behavior or finding fault with their response to the gradually unfolding evidence of fraud, especially since guidelines for NIH

*The outside investigators appointed by the NIH were Howard E. Morgan, MD, Evan Pugh Professor of Physiology at the Pennsylvania State University College of Medicine, Hershey, Pa, chairman; Francis J. Klocke, MD, chief of cardiology and professor of medicine and physiology, State University of New York at Buffalo; John T. Shepherd, MD, dean, Mayo Medical School, Rochester, Minn; and C. Kern Wildenthal, MD, dean of Southwestern Medical School at the University of Texas Health Science Center at Dallas.

grantees in such situations were (and still are) lacking. But the NIH stood by its characterization of events, merely releasing copies of the Harvard researchers' rebuttals to journalists who requested them, along with the official report.

Braunwald's position is that Darsee was so clever and persuasive a cheater that it took almost a year to "crack the case" and determine the true extent of his fraud. In the first instance, confronted with only a single acknowledged act of fabrication, Braunwald explains, he was unwilling to ruin his protégé's promising research career, although he promptly terminated Darsee's NIH fellowship—without telling the NIH why—and told him he was through at Harvard. Still, he kept Darsee on at the laboratory throughout the summer and early autumn of 1981 to carry on work with the NHLBI contract and finish other projects.

Braunwald denies that Darsee's extensive cheating is symptomatic of any general problem with the administration of his laboratory—though a number of administrative reforms were implemented after Darsee's misbehavior was discovered. In an interview following the release of the NIH report, Braunwald complained bitterly that he and Kloner are being blamed for permitting the fabrication to happen under their auspices. "We cracked the case and, not only that, we provided the evidence that he cheated in other places," Braunwald said. "I think it's a bum rap to cast all the aspersions at us. I'm not saying we're perfect, but we did catch him."

'A Legendary Figure'

By all accounts, Darsee was an unlikely cheat who would have gone far on his more honorable talents. Born in West Virginia in modest circumstances, he worked his way through the University of Notre Dame and the Indiana University School of Medicine and went on to become "a legendary figure," as one Emory professor put it, during his postgraduate training in Atlanta. He was recommended to Braunwald in 1978 in superlative terms: "He stands head and shoulders over many compatriots, yet is extremely well liked and respected," wrote Nanette K. Wenger, MD, professor of cardiology at Emory. "He has an unusually fine blend of intelligence, self-discipline, ambition, and personality," added Paul F. Walter, MD, chief of the coronary care unit at the Veterans Administration Hospital in Decatur, Ga. "I haven't recommended a candidate so highly in years."

Darsee came to Harvard at the age of 31 with dozens of published papers and abstracts in his bibliography. He proceeded to dazzle his new Boston colleagues with his adeptness in the laboratory, his creativity, and above all his awesome productivity. "He was a hard worker and a very efficient worker—one of those people who always has a clean desk," recalls one of his Harvard co-workers. "He worked like a machine." Adds another: "His work was good, he was intelligent, he wrote well, he was the best public speaker the Brigham had seen in a long time."

Braunwald often called Darsee one of the most outstanding of the 130 research fellows he has trained, more than 40 of whom are now full professors, department chiefs, or directors of academic cardiology divisions. Robert I. Levy, dean of Tufts University Medical School, Boston, former NHLBI director, and a member of the Harvard blue ribbon panel appointed by Tosteson to investigate Darsec's work, commented in a recent interview with *JAMA* MEDICAL NEWS: "Everything Darsee did was two or three standard deviations above the norm."

In retrospect, of course, such confidence was undoubtedly misplaced. And new disclosures indicate that Darsee may have started cheating a long time ago. In a recent letter to Braunwald, whom he does not know, Julian R. Pleasants, PhD, associate professor of microbiology at Notre Dame, recalled two articles by Darsee that were published in 1969 in the university's undergraduate science magazine, on which Darsee was a staff member.

In one of the reports, Darsee claimed to have done assays of ten different hormone levels twice weekly on 50 rats beginning at 1 week of age and continuing to death. Allegedly, he found that six of the hormones declined with age of the animals. He then claimed to have taken 100 older rats and injected them thrice weekly with the hormones whose levels he had found to decline, thereby bringing the quantities up to juvenile levels. By this means, he said, he succeeded in extending the average life span of the treated animals by 47% in comparison with the average life span of 50 control rats.

All data were presented as Darsee's own, Pleasants pointed out; no faculty member in whose laboratory the work had been performed was listed. Now in such articles, the name of a faculty member must be listed.

"Such experiments would be extremely demanding," Pleasants said in a recent interview with *JAMA* MEDICAL NEWS. "Among other things, it is very difficult to bleed a young rat from a tail vein twice a week and to inject precise amounts of hormones into rats' femoral veins three times a week. The animals involved would be in a constant state of stress and have continual trauma, fibrosis, and infection. Thus, it is quite implausible for anyone, much less an undergraduate working in his spare time, to accomplish such experiments."

Darsee also advertised and gave a seminar on these and other hormone experiments. Asked why no one questioned the work at the time, Pleasants replied, "When I heard it, I thought he was summarizing the work of a whole laboratory over a considerable period of time. It just slipped by."

Indeed, it was Darsee's too-good-to-be-true performance that eventually brought about his downfall,

continued on page 1802

continued from page 1799
though it took a long time for Braunwald and Kloner to accept the full truth about their young protégé. By the spring of 1981, several of his co-workers had concluded independently that he was a fraud. "It was the volume and the cleanness of the data that couldn't be explained," said one person who was familiar with the evolving suspicions in Braunwald's laboratory. What aroused doubts in particular was Darsee's claim to have done some complicated long-term dog experiments, trials spanning up to six weeks and involving indwelling intracoronary pressure transducers, ultrasonic crystals, and meticulous sterile techniques. No one in the laboratory could remember seeing him do such arduous experiments in the numbers he claimed. The suspicions first surfaced at a data meeting in March 1981, but no one felt the evidence was solid enough then to make a direct accusation.

After that meeting, several of the other research fellows and technicians began comparing notes and voicing their suspicions about Darsee. Though even now no one will be quoted by name, direct sources told *JAMA* MEDICAL NEWS that they "cautiously" agreed among themselves to keep a close watch to develop the kind of evidence that would enable them to go to Kloner, who seemed oblivious to the possibility that Darsee might be "manufacturing" data. (Kloner has rebuffed repeated attempts to interview him on the subject.)

Several Days in May

According to the chronology put together by the NHLBI investigation, things finally came to a head on May 14, 1981. On that day, Darsee was observed doing an experiment that normally would involve dissecting a dog's heart to determine coronary blood flow on the basis of the distribution of injectable radioactive microspheres. At the end of the day, Darsee was seen placing an animal, dog D-35, in a disposal bag with its heart intact. Sharon Hale, the laboratory's chief technician, and Edward J. Brown, MD, another research fellow

in the laboratory who has since joined the faculty of the State University of New York at Stony Brook, decided to remove the heart from dog D-35 and retain it. This was "because of suspicions," the NHLBI report said, "that there were instances in which Dr Darsee was reporting myocardial blood flow data on hearts that had not been analyzed."

The next day Hale and Brown went to Kloner to voice their conviction, based on circumstantial evidence, that Darsee was an habitual fraud. Inexplicably, they did not mention the incident with dog D-35 the day before, Braunwald maintains. Interviewing others, Kloner found the suspicions were shared. Investigating further, he came on a photocopy of an undated strip chart from one of Darsee's studies that purported to show hemodynamic and other cardiac data on a single animal, without interruption, over 24, 48, and 72 hours and at seven days. He concluded, in the words of the NHLBI investigators, that it was "highly unlikely" that the recording apparatus had not been used for any other studies during that entire week.

Kloner summoned Darsee and asked to see all primary tracings for the same study, which Darsee was planning to submit to the American Heart Association (AHA) for presentation at its November 1981 meeting in Dallas.

It is not clear why Kloner did not demand to see the raw data immediately. What is known is that six days later, on May 21, several colleagues looked on in amazement as Darsee put a greyhound on the table, hooked it up to various monitoring instruments, ligated a coronary artery and proceeded to obtain successive tracings of intraventricular and arterial pressures, myocardial segment length, ECG and other parameters, turning the dials on the machine and altering the scales to reflect ostensible infarct size at different times. He labeled the tracings, obtained in one afternoon, as "24 hours," "72 hours," "one week," and "two weeks." Then he carefully cut the borders from the strip charts to make it appear that the tracings were not continuous.

One of the research fellows called Kloner, who was at the Brigham and Women's Hospital a block away from the Cardiac Research Laboratory, and said: "Come over, Bob. We'll give you something that will make you certain."

Confronted the next day by Kloner, who had fished the cut-out borders from the trash, Darsee admitted what he had done, but insisted that it was the only time he had ever fabricated data. Over the long Memorial Day weekend, Braunwald pondered what to do. On Tuesday, May 26, he interviewed the other fellows and the technician, then called Darsee in and told him the proffered assistant professorship at Harvard, along with the directorship of his own laboratory at the Beth Israel Hospital, were being withdrawn. Braunwald asked Darsee to stay on at the laboratory for a few more months to help with an investigation of his work and wrap up unfinished business—including the federally funded AMPIM study. But during that period he would be paid by departmental funds, not NIH fellowship money.

Later Braunwald explained that he elected not to inform the NIH of the episode because he considered it a "single, bizarre act" and because there were no NIH guidelines either requiring disclosure or ensuring that disclosure would remain confidential. Similarly, he did not inform Emory colleagues, who that very week were sending Braunwald enthusiastic recommendations for Darsee's elevation to the Harvard faculty. To reveal such information to outsiders, Braunwald said in an interview in March 1982, would have been "the academic equivalent of capital punishment. I did not want to take that on myself. I did not want to foreclose a career in academic medicine for such a gifted person on the basis of what appeared to me to be bizarre behavior. It would have been a tragic loss."

There were at least two other factors in Braunwald's decision. The thought occurred to him that if he let outsiders know there was a problem with Darsee's integrity, "I would be leaving myself wide open to a lawsuit." In addition, he believed that the strong suspicions about Darsee were motivated in large part by "jealousy" among the other fellows. "It's fair to say there was a lot of jealousy," Braunwald said, though he acknowledged that technician Hale "would have no motive."

Asked recently why he did not consider the first suspicious strip chart, found by Kloner before May 21, as evidence of a possible pattern of fabrication, Braunwald replied: "I viewed that as just part of a single major deception."

According to various sources, Darsee lashed out at his accusers in an attempt to shift the burden of proof to them. In a letter to Braunwald written six months later, Darsee said that his behavior should be considered in the context of the "laboratory environment," which he characterized as bristling with envy, spite, spying, and even sabotage directed toward him—especially by Ed Brown—because Darsee was such a hard worker, meticulous scientist, and lone wolf. There was "enormous resentment toward me," Darsee wrote.

Braunwald and Kloner commenced what the senior investigator later called a "hard look" at Darsee's work, published and unpublished. Comparing Darsee's findings against others' work, and sometimes repeating experiments, "everything checked out," Braunwald said. Although Darsee could not produce large portions of his raw data, there was a disposition to believe the "single bizarre act" hypothesis. Consequently, Braunwald saw no "big risk" in permitting the submission of numerous papers and abstracts on which Darsee was principal author or coauthor. In August 1981, for instance, Braunwald (the only cardi-

continued on next page

continued from previous page
ologist member of the National Academy of Sciences) submitted to the *Proceedings of the National Academy of Sciences* (*PNAS*) a paper entitled "Recovery From Prolonged Abnormalities of Canine Myocardium Salvaged From Ischemic Necrosis by Coronary Reperfusion." Darsee was one of five coauthors. Criticized for this later, Braunwald said that the *PNAS* paper had "passed special scrutiny."

"I really believed in this man," Braunwald subsequently explained. "I really fell into the trap of thinking this was one bizarre act."

Braunwald's belief in Darsee was seriously shaken, however, by the events of October 1981, when the NHLBI distributed unblinded data from the four AMPIM study centers that were working on dog models. On Oct 6, NHLBI staff called the four centers to inform them of an unexplained discrepancy between the end-point data (the ratio of infarct size to total left ventricular size) from the Brigham and the other centers. The Brigham's data showed particularly large favorable effects of ibuprofen and verapamil on infarct size. By Oct 21, it was apparent that the anomalies extended to other data as well; the Brigham's data on heart weight and myocardial blood flow were also noncomparable. Another week of close analysis by NHLBI staff revealed other curiosities.

Meanwhile, in Boston, Kloner was doing his own analysis. Graphing Darsee's pre-May 22 data against his post-May 22 data, Kloner discovered a striking difference. The pre-May 22 blood flow measurements showed a standard deviation (SD) of only 0.04 to 0.05, compared with SDs in the other centers ranging from 0.25 to 0.33. After May 22, Darsee's SDs were 0.31. The mean heart weight in Darsee's dogs was far in excess of those reported elsewhere, but that, too, came into line after the May 22 watershed.

Asked to produce the raw data for his pre-May 22 AMPIM experiments, Darsee could not. He explained the discrepancies between his and other investigators' data by claiming superior technique. Kloner called the NHLBI on Oct 29 to suggest that Darsee might be the solution to the AMPIM mystery, a revelation he repeated the next day at a Bethesda meeting of all the co-investigators. NHLBI staffers were incensed that they had been given no earlier clue and that Darsee had been permitted to work on the project after doubts of his integrity arose.

More surprises were in store. Leafing through the annual meeting issue of *Circulation* in early November 1981, Braunwald says he was startled to find Darsee's name on four abstracts from Emory of which he had no prior knowledge, plus two others from the Brigham in collaboration with radiologist Leonard Holman, MD. The latter compared assessment of myocardial infarct size by creatine kinase levels and by photon emission tomography. Examining the Darsee-Holman abstracts, Braunwald determined that the correlations between the two methods were "unbelievably good." He called Holman, who was flabbergasted to learn that Darsee's work was under suspicion and that he was not going to head up the Beth Israel cardiac research laboratory, as Darsee had led Holman to believe. Darsee, who had nine abstracts listed on the AHA meeting program that year—more than any other researcher—was told he could not present the Harvard work. He did present the Emory papers, however, and Braunwald did not inform the Emory investigators of his discoveries about Darsee until the actual meeting in Dallas.

The Harvard Investigation

With this expanding and irrefutable evidence of systematic fraud on their hands, Braunwald and Tosteson decided in November that it was time to appoint an impartial investigative committee. Tosteson picked eight members, including three from outside Harvard; the panel was chaired by Richard S. Ross, MD, dean of the Johns Hopkins School of Medicine. The committee met only twice, in December 1981 and in January 1982. Dean Levy of Tufts recalls Darsee from his appearance before the group as a "charming young man" who held firmly to his insistence that the only wrongdoing had been the May 21 fabrication. "When we asked him how he could have reported all the data from Emory two years later if he had not learned to keep raw data," Levy recalled recently, "he had no answer except to smile."

Though the panel asserted in its Jan 25, 1982, report that it had interviewed or received written testimony from "all the principals," it never interviewed the laboratory co-workers who blew the whistle on Darsee. This omission greatly disturbed the whistleblowers and reportedly astounded Harvard Medical School Associate Dean S. James Adelstein, MD, PhD, who was asked a year later to conduct another internal review, the results of which have never been made public. Levy said the panel was "made aware that some investigators suspected John Darsee had been cutting corners for some time. But our charge was not to define how much misbehavior had occurred but whether Eugene Braunwald had acted appropriately and effectively."

Nevertheless, the Ross committee declared it was convinced that the extent of the fraud had been established by Braunwald and Kloner, whose audit it accepted unquestioningly. A year later, when it appeared that the fraud was considerably more widespread, Tosteson reportedly wrote the NIH to retract that sentence from the Ross committee's report, though Harvard made no public announcement of the revision.

One incident from late 1981 suggests that Braunwald was still disposed to be skeptical of Darsee's

detractors, even though events were fast proving them correct. In December, when the Ross committee was beginning its work, laboratory staffers got a look at Darsee's notebooks and discovered that he had entered blood flow calculations and gamma counter printout data (reflecting the distribution of circulating radiolabeled microspheres) for dog D-35, the animal that he had discarded the previous May without removing its heart. These calculations would require not only removal of the heart but its dissection.

Braunwald was informed of this further piece of evidence, which pointed to a wider pattern of fabrication than the audit had yet turned up. But he did not mention it to the Ross committee and he doubted the veracity of the evidence because there was no proof that the intact heart that he at one point was shown was actually from dog D-35. In other words, he again doubted the accusers. "There was enough of a cloud about his relationship to Brown and a question of whether Brown was 'out to get him' that I felt we had much more serious stuff on him than dog D-35," Braunwald said recently. "Maybe in retrospect we should have made a bigger deal of it. But we're not trained as sleuths."

The Ross committee's report and the Braunwald-Kloner findings on which it relied came in for some of the NHLBI panel's harshest criticism. While approving two of the Ross panel's major recommendations—that a committee of uninvolved professors should have been appointed soon after the May incident, and that a systematic search should have been made to identify other Darsee collaborators at the school and inform them of the suspicions—the NHLBI group pointedly concluded: "It would have been preferable for Harvard officials to have insisted on an exhaustive report without unrealistic time constraints." The rushed pace of the Ross committee's work "hampered an in-depth investigation that might have led to discovery of the full extent of data irregularities," the federal inquiry said. "It would have also been preferable for the committee not to have relied so heavily on the internal audit conducted by Drs Braunwald and Kloner," which the NHLBI panel found "insufficiently rigorous." And it added that Braunwald should have told the Ross committee about dog D-35, even if the provenance of the preserved heart could not be established.

Especially in light of the absence of much of Darsee's primary data and misleading descriptions of the randomization of experimental animals, the NHLBI group said, Braunwald and Kloner should have suspected a wider pattern of fraud at a much earlier date. Furthermore, they should have listened more seriously to Darsee's co-workers: "Not all relevant allegations by co-workers in the laboratory were adequately investigated in 1981," the NHLBI panel said.

In any case, the Ross committee's report in early 1982 failed to have its intended effect, which was to lay the Darsee matter to rest. The NHLBI had by then decided on its own investigation, and many in the scientific community were expressing private skepticism that a man whose fraud encompassed one major NIH study as well as three other studies did not, as the Ross panel asserted, perpetrate other misdeeds.

Consequently, Braunwald and Kloner renewed their efforts to review the corpus of Darsee's work. By this time, they had learned to "think like sleuths." Kloner hit on the most telling test. He and Braunwald had noticed discrepancies between Darsee's notebooks and records in the logbook for disposal of radioactive dogs. If Darsee had actually performed blood flow studies on his dogs, the heart tissue would have to retain residual radioactivity and microscopic evidence of microspheres. Analyzing tissue blocks from seven pre-May 22 dogs from the AMPIM study, Kloner found no evidence of radioactivity or microspheres. Seven of the 15 dogs Darsee used after May 22, however, did show residual radioactivity and microspheres. Tests of the few available tissue blocks from Darsee's other studies showed the same absence of evidence that the requisite tests were done. The results of Kloner's review were

continued on next page

continued from previous page
confirmed at the NIH and at another institution.

Largely because of this clinching evidence, Braunwald and Kloner were forced in July 1982 to write letters of retraction for all or portions of nine papers coauthored with Darsee, including the *PNAS* paper that Braunwald had earlier insisted had received special scrutiny.

Braunwald, who failed earlier to get the NHLBI panel to expand its scope to Darsee's work at Emory, says he has found a number of problems with those papers. For instance, he says that two papers that list Darsee as first author, published in 1980 in the *American Journal of Cardiology*, show "a much [unbelievably] closer correspondence between measurements of cardiac output between two relatively inaccurate methods than would be anticipated from duplicate measurements using either method alone." Another paper, of which Darsee also is first author, published in the *New England Journal of Medicine* in 1981, presents results on complete 24-hour urine collections in every member of an extended family of 46 persons, including an infant—a fairly remarkable achievement. That paper, according to Braunwald, also contains anomalous data on creatine clearance. Braunwald says that Emory's investigation also has turned up evidence that, among other things, some of Darsee's coauthors "probably never witnessed Dr Darsee performing the reported experiments or saw the data he purportedly collected," that some may not have been aware he was listing them as coauthors, and that he may have acknowledged some individuals who "may not even have existed either at Emory or at Harvard."

Why? Why?

Amid the debris of the Darsee affair, as with a number of scientific fraud cases in the past decade, the question lingers: What drove him to it? The mystery is doubtless unresolvable, perhaps even for the perpetrator, but the Darsee affair has certainly occasioned more explicit examination of the issue than many—but not all—preceding cases (see "The Sloan-Kettering Affair," *JAMA* [MEDICAL NEWS] 1974; 229:1391-1410). After all, Darsee apparently would have had a brilliant career without cheating. Furthermore, though his direct testimony is conspicuously lacking, the statements he is known to have made do not suggest that he has faced up to his deeds. First there was his letter to the NIH, in which he said he had no recollection of fabricating any data apart from the May 21 incident but asked forgiveness "for whatever I have done wrong" nonetheless.

Then, in a December 1981 letter to Braunwald, Darsee implies he was framed by envious colleagues and excuses the one instance of admitted fabrication by the enormous stress he was under: "This was an extremely difficult period for me," he wrote. "I had too much to do, too little time to do it in, and was greatly fatigued mentally and almost childlike emotionally. I had not taken a vacation, sick day, or even a day off from work for six years. I had put myself on a track that I hoped would allow me to have a wonderful academic job and I knew I had to work very hard for it. Successful people don't take time off—at least that's the premise on which I based my work."

The Darsee affair also poses more squarely than most other cases of research fraud whether something in the highly competitive research enterprise, some flaw in the system, pushes talented scientists to cheat, gives them too much opportunity to do so, and makes it fiendishly difficult to uncover the extent of it when it occurs. "Although these issues have been raised before," NIH Associate Director William F. Raub, PhD, said in a recent telephone interview, "the Darsee case raises them most explicitly." Braunwald himself, though he disclaims responsibility for contributing to Darsee's transgressions, mused in a letter to Emory's Chief of Medicine Willis Hurst, MD, in December 1981: "We must ask ourselves if there is something inherently wrong in the system, the person, or both, that allows such behavior to emerge."

"I don't go along with the idea that if you're under pressure, it's an excuse for cheating," Thomas C. Chalmers, MD, dean of Mount Sinai School of Medicine in New York, said in a recent interview with *JAMA* MEDICAL NEWS. "Cheating is not just an abrupt and isolated phenomenon in research, it's on the end of a slope. All the studies on medical students have revealed shockingly high rates of cheating to get into medical school or in medical school. The students who are cheating aren't the dumb ones. They're often the brightest. We have created a milieu in which absolute honesty is not the be-all and end-all."

Like a number of medical schools these days, Mount Sinai—which last year had its own case of research fraud to untangle—is in the process of writing guidelines for dealing with suspected misconduct. Similar in some respects to the suggested guidelines published last year by the Association of American Medical Colleges, the Mount Sinai guidelines will go beyond post hoc procedures to suggest prophylactic measures. According to Chalmers, the guidelines will stipulate "that research should never be solitary, that co-investigators be responsible for all the data, and that no data be presented publicly that hasn't been presented in-house for critical review. And the busy lab chief shouldn't be so busy that he doesn't have the time to look at the original data and verify it to his satisfaction."

Harvard also is considering new guidelines for dealing with suspected scientific misconduct, but they are focused on post hoc procedures rather than on preventive measures.

The Darsee affair is an "extreme case," and has "raised everybody's consciousness and led us to enunciate guidelines and procedures that should have

been in place before," says Levy of Tufts. "We're entering a new era of consciousness because we need to ensure the integrity of our product. But will the sensitivity be there three years from now? You pass out too many written words and people just put them on the shelf." Adds Braunwald: "We need to be more vigilant. But we have to be careful we don't develop an atmosphere of repression."

This leads to another aspect of such cases: that when, even early in a budding scientist's career, a technician, teacher, laboratory chief, or peer reviewer perceives something as "not quite right," that that situation be confronted—however delicately—at the time. The possible result might be that instead of passing a cheater on to another laboratory—as is all too commonly done—the person's career in scientific fraud would be cut short.

Indeed, perhaps the most disturbing postscript on the Darsee affair comes from a Harvard researcher who was knowledgeable about the details of the unfolding saga. "The man could have gently faked his way through a lifetime of research and have been greatly honored," said this scientist, who preferred anonymity. "We all know scientists like that."

—by RICHARD KNOX

Emory Reports on Darsee's Fraud

John R. Darsee, the young researcher who was forced to leave Harvard Medical School nearly 2 years ago after he was caught faking data, apparently has had a long and prolific career in data fabrication. Prior to going to Harvard as a fellow in Eugene Braunwald's cardiology laboratory at the Brigham and Women's Hospital, Darsee was in training at Emory University from 1974 to 1979. At least ten full-length papers and 45 abstracts, which Darsee coauthored with some of Emory's most prominent researchers, have been published on the basis of his work there. Most of them have now been retracted after an Emory investigation concluded that Darsee's work cannot be validated.

The Darsee case, which has raised serious questions about the way in which young investigators are supervised during their training, also raises issues about coauthors who allow their names on publications even though they have not done the research. The latter phenomenon has been revealed with particular clarity as the result of the Emory investigation. In several cases, senior scientists, in response to queries from the investigating committee, explained their failure to detect data fabrication by denying any knowledge of the research on which Darsee reported. Darsee, they assert, simply put their names on abstracts without their knowledge, and none of the Emory scientists apparently insisted that his or her name be removed.

According to Emory officials, only two of the ten papers are still considered valid. Of the 45 abstracts, again only two stand up to scrutiny. In a 5 May statement, Neil C. Moran, chairman of the Emory committee that spent more than a year on the Darsee case, said, "Many of the abstracts appear to be completely fictitious."

Donald O. Nutter, executive associate dean of the medical school, was among the senior faculty whose names Darsee added to published abstracts. In a telephone interview with *Science*, Nutter said, "I told John it was inappropriate behavior." But Nutter failed to take the extra step of insisting that his name be removed because Darsee was "fluent and convincing" when he argued that he owed Nutter credit because of his role in Darsee's training.

Virtually everyone who came in contact with Darsee at Emory (and later at Harvard) found him to be exceptionally able. Senior faculty quickly entrusted a lot to him. The one person who ever questioned the integrity of Darsee's work at Emory was Ayten Someren, a pathologist who reported data discrepancies to his superiors. But no substantial effort to investigate was made, and there was no mention of it in letters of recommendation Emory faculty sent to Harvard.

Moran, investigating chairman, reports that his committee did not go beyond its charge, which was to determine the extent of Darsee's data fabrication, to an analysis of how such widespread fraud could go undetected. (The Emory investigation took place only after Braunwald told researchers there that he was suspicious of material in Darsee's papers, much of which he found not scientifically creditable.)

A full study of the "why and how" of the Darsee case, Moran said, would have required a more lengthy, complex investigation than the committee was prepared to undertake. However, he said, "How it happened in one way is quite simple. When all principles involved neglect to take full intellectual responsibility for the research, including examination of data, statistical analysis, and assessment of expert consultation, it is easy for a person inclined to dishonesty to get away with it."

In one instance of reputed fraud, Darsee coauthored a paper with cardiologist S. B. Hyemsfield in *The New England Journal of Medicine* on taurine levels in heart patients. The Moran committee report states that "there are no original records of this project at Emory University." Heymsfield does not remember the patients. The published paper does not identify any hospital or clinic as the place the patients came from. Furthermore, the acknowledgment at the end of the paper expresses "indebtedness" to three scientists who apparently do not exist. A Johnson S. Caulder, Ph.D., is, for instance, credited with doing amino acid analyses for the study. Yet no such person has ever worked at either Emory or Harvard. Moran thinks this should have been a clue to Darsee's supervisors. "It is not unusual to have interlaboratory collaboration," Moran says, but the senior investigators ought to actually know their collaborators.

Darsee did not meet personally with the Moran committee, but he did respond by letter to some of its inquiries. " 'The names of the three persons acknowledged at the end of the Taurine paper were mentioned in the hospital chart of one of the family members I first saw in clinic when I was an intern in 1974,' " the report quotes as Darsee's reply. " '. . . . Neither Steve Heymsfield nor I had any correspondence with them. They are not from Georgia but I don't remember where any of them are from.' "

At present, Darsee, who declines to talk to the press, is working at Ellis Hospital in Schenectady, New York, where he has recently been shifted from a training program in critical-care medicine to working with computers. Darsee has been "removed from patient contact," according to hospital spokesman Andrew Foster who says that the hospital, which has stood by Darsee, is "trying to ascertain his place in medicine." Hospital staff, Foster says, frequently describe Darsee as being "brilliant" and "charismatic."

Emory promptly sent a copy of the Moran report to the National Institutes of Health but was troubled by the issue of sending it directly to Ellis Hospital administrators or medical licensing boards. "We have come to no conclusion and no consensus about our responsibility here," Moran says, noting liability as one concern. "One argument against direct notification is that we would be questioning his competence as a physician and we have no basis for that. What we have is evidence of dishonesty in research." But, Moran says, "I'm not trying to defend our lack of action." According to Ellis Hospital spokesman Foster, the hospital has requested a copy of the report and has been told one will be sent.

Although there has been no formal faculty discussion of the aftermath of the Darsee case at Emory as yet, the subject will come up in June at a meeting of department chairmen. Much of the response to date, Moran reports, has been what he calls the "bad apple, these things don't happen very often sort." Says he, "I don't think that's adequate."—BARBARA J. CULLITON

RETRACTION

Darsee JR, Heymsfield SB, Nutter DO. Hypertrophic cardiomyopathy and human leukocyte antigen linkage: differentiation of two forms of hypertrophic cardiomyopathy. N Engl J Med 1979; 300:877-82.

To the Editor: In response to the announcement by Harvard University that John R. Darsee, M.D., had fabricated research data while serving as a research fellow in their school of medicine, we initiated an extensive investigation by both internal and external committees of Dr. Darsee's work at Emory University. This letter concerns a publication in the *Journal* that was authored by Dr. Darsee (Darsee JR, Heymsfield SB, Nutter DO. Hypertrophic cardiomyopathy and human leukocyte antigen linkage: differentiation of two forms of hypertrophic cardiomyopathy. N Engl J Med 1979; 300:877-82).

The evidence available for review by our investigating committees consisted of incomplete laboratory records and clinical records, and this material failed to sustain the results and conclusions presented in the article. This general observation is supported by the following specific points: *(1)* Patients with coronary-artery disease were included in the proband group despite the disclaimer in the Methods section of the paper. *(2)* We cannot find laboratory reports for the 70 matched controls. *(3)* The diagnosis of hypertrophic cardiomyopathy cannot be confirmed from the hospital records of at least six of the probands. One of these individuals is the proband for the family reported in Figure 3 of the paper. *(4)* The reported correlations between laboratory and clinical data cannot be confirmed. The paper reports that 11 of 12 white probands who were B12 antigen-negative had hypertension. The hospital records indicate that at least six were normotensive. The paper reports that only 3 of 33 black probands positive for the B5 antigen complex had hypertension. The hospital records indicate that at least 10 of these patients were hypertensive. *(5)* Nine family pedigrees were either published in the article or were contained in the NAPS document. However, laboratory data are available for only 21 of the 86 members of these families. Most distressing is the fact that no laboratory data are available for the largest family (Fig. 1 in the *Journal* article), and data for the proband of this family are not included in the NAPS listing of 22 white probands. The proband for the family presented in Figure 2 is described as being hypertensive, but this cannot be confirmed from the hospital record. Finally, as previously noted the hospital record does not confirm the diagnosis of hypertrophic cardiomyopathy in the proband of Figure 3, and furthermore this patient is described as being normotensive, but the hospital record indicates the presence of hypertensive cardiovascular disease.

In view of this information we believe that the paper must be considered invalid and should be retracted.

We apologize to you and the *Journal* for any embarrassment or inconvenience that may result from our investigation of Dr. Darsee's research activities and publications.

Donald O. Nutter, M.D.
Professor of Medicine (Cardiology)

Steven B. Heymsfield, M.D.
Associate Professor of Medicine

James F. Glenn, M.D., Dean
Emory University
Atlanta, GA 30322 School of Medicine

RETRACTION

Darsee JR, Heymsfield SB. Decreased myocardial taurine levels and hypertaurinuria in a kindred with mitral-valve prolapse and congestive cardiomyopathy. N Engl J Med 1981; 304:129-35.

To the Editor: In a recent disclosure, Harvard University announced that a research fellow in the school of medicine, Dr. John R. Darsee, had fabricated data that formed the basis of several publications. Prior to Dr. Darsee's fellowship, he was a resident in the Department of Medicine at the Emory University School of Medicine. Subsequent to the Harvard report, Emory University had an internal and external committee review Dr. Darsee's publications during his years at Emory. This letter concerns a publication in the *Journal* that was authored by Dr. Darsee (Darsee JR, Heymsfield SB. Decreased myocardial taurine levels and hypertaurinuria in a kindred with mitral-valve prolapse and congestive cardiomyopathy. N Engl J Med 1981; 304:129-35).

Our recently completed investigation disclosed that *(1)* Dr. Darsee's laboratory notebook had only patients' initials, and therefore cross-checking results with hospital charts was not possible; *(2)* it is improbable that Dr. Darsee could have obtained fresh myocardial tissue within four hours of death in 25 subjects; *(3)* it is highly unlikely that venous blood and 24-hour urine specimens could have been reliably collected twice a year from all 43 persons, particularly in the two-year-old child; and *(4)* the referring physicians and collaborators acknowledged at the end of the paper could not be identified.

In view of our inability to verify the observations in this paper, we believe that the results should be considered of uncertain validity. We therefore wish to retract this publication. We apologize to you and the *Journal* for any embarrassment or inconvenience that may result from our investigation of Dr. Darsee's research activities and publications.

Steven B. Heymsfield, M.D.
Associate Professor of Medicine

James F. Glenn, M.D., Dean
Emory University
Atlanta, GA 30322 School of Medicine

The New England Journal of Medicine

Official Organ of
The Massachusetts Medical Society

Goodwill M. Stewart, M.D.
President

William B. Munier, M.D.
Executive Vice-president

Charles S. Amorosino, Jr.
Executive Secretary

PUBLISHED WEEKLY BY THE COMMITTEE ON PUBLICATIONS OF THE MASSACHUSETTS MEDICAL SOCIETY

Samuel K. Stewart, M.D., *Chairman*
John C. Ayres, M.D.
Frank E. Bixby, Jr., M.D.
John I. Sandson, M.D.
William H. Sweet, M.D., D.Sc.
Robert E. Tranquada, M.D.
Percy W. Wadman, M.D.

Arnold S. Relman, M.D., EDITOR
Marcia Angell, M.D., DEPUTY EDITOR
Edwin W. Salzman, M.D., DEPUTY EDITOR

ASSOCIATE EDITORS

Jane F. Desforges, M.D.
Ronald A. Malt, M.D.
Norman K. Hollenberg, M.D., Ph.D.
Morton N. Swartz, M.D.
Franklin H. Epstein, M.D.

Francis D. Moore, M.D., BOOK REVIEW EDITOR
John C. Bailar, III, M.D., STATISTICAL CONSULTANT
John K. Iglehart, SPECIAL CORRESPONDENT

Joseph J. Elia, Jr., MANAGER OF EDITORIAL OPERATIONS
Emily S. Boro, DIRECTOR OF COPY EDITING
Marlene A. Thayer, EDITORIAL OFFICE MANAGER

EDITORIAL BOARD

Richard H. Egdahl, M.D.
Robert J. Mayer, M.D.
Kenneth J. Rothman, Dr.P.H.
Kurt J. Bloch, M.D.
Paul Calabresi, M.D.
Aram V. Chobanian, M.D.
John T. Harrington, M.D.
Homayoun Kazemi, M.D.
Kenneth McIntosh, M.D.
David G. Nathan, M.D.
Lawrence G. Raisz, M.D.
Thomas J. Ryan, M.D.
Samuel A. Latt, M.D., Ph.D.

Frederick Bowes, III, DIRECTOR OF PUBLISHING OPERATIONS
Ronald H. Brown, MANAGER OF ADVERTISING & MARKETING
William H. Paige, MANAGER OF PRODUCTION & DISTRIBUTION
Milton C. Paige, Jr., CONSULTANT

PROSPECTIVE authors should consult "Information for Authors," which appears in the first issue of every volume and may be obtained from the *Journal* office.

ARTICLES with original material are accepted for consideration with the understanding that, except for abstracts, no part of the data has been published, or will be submitted for publication elsewhere, before appearing in this *Journal*.

MATERIAL printed in the *New England Journal of Medicine* is covered by copyright. The *Journal* does not hold itself responsible for statements made by any contributor.

NOTICES should be sent at least 30 days before publication date.

ALTHOUGH all advertising material accepted is expected to conform to ethical medical standards, acceptance does not imply endorsement by the *Journal*.

REPRINTS: The *Journal* does not stock reprints, and reprints of the MGH CPCs are not available.

SUBSCRIPTION PRICES: USA: $52 per year (interns, residents $35 per year, students $30 per year). Canada (U.S. funds only): $62 per year (interns, residents $45 per year, students $40 per year). Mail checks to Subscription Payments, P.O. Box 4772, Boston, MA 02212.

EDITORIAL OFFICES: 10 Shattuck St., Boston, MA 02115.

BUSINESS AND SUBSCRIPTION OFFICES: 1440 Main St., Waltham, MA 02254.

LESSONS FROM THE DARSEE AFFAIR

It seems paradoxical that scientific research, in many ways one of the most questioning and skeptical of human activities, should be dependent on personal trust. But the fact is that without trust the research enterprise could not function. This trust is manifested in many ways. Scientists habitually challenge the validity and the interpretation of the experimental evidence reported by their colleagues at meetings or in the literature, but they trustingly assume that the evidence has been honestly gathered and reported. In their own work they are often dependent on raw data obtained by technicians, fellows, or collaborators. They may personally review the data and check the calculations, but unless they have made the primary observations themselves they must rely on the honesty of those who have. Likewise, editors and referees of scientific papers, even while they search for possible errors in the manuscripts they review, have no choice but to assume that the authors have honestly reported what they did and what they observed.

There is no practical alternative to this presumption of honesty in research, because it would be impossible to verify every primary datum and every descriptive statement in a research report, just as it would be impractical for senior investigators to check every primary observation made by their collaborators and assistants. Furthermore, any attempt to police the system would probably create such an atmosphere of suspicion and distrust that the free spirit of scientific inquiry would be crushed. Research is a collegial activity that requires its practitioners to trust the integrity of their colleagues.

Violations of that trust are probably not as common as the publicity they receive suggests, but whatever their frequency, they are always a reason for serious concern and soul-searching because they threaten the very foundations of research. A case in point is the truncated research career of Dr. John R. Darsee, a young clinical investigator who was recently found to have fabricated an extraordinary series of published findings, beginning when he was an undergraduate biology student at Notre Dame, continuing through his medical residency and cardiology fellowship at Emory University in Atlanta, and ending with a fellowship in the cardiac research laboratory of the Brigham and Women's Hospital, a Harvard teaching affiliate in Boston. The long and complicated story of how Darsee was first caught fabricating data by his associates and supervisors at Harvard in May 1981, and of how that revelation ultimately led to the unraveling of his career of fraud and deception, need not be repeated here, because it has been in the headlines many times during the past year. A particularly good summary was published in *Science* a few months ago,[1] and other news articles continue to appear.

While he was at Emory, Darsee's fabrications compromised the integrity of at least eight published papers, which are now being retracted. In addition, he

appears to have manipulated or invented the data published in at least 32 abstracts, which are also being retracted. Most of these articles and abstracts had Emory faculty members as coauthors, but according to the recently released findings of an ad hoc review committee at Emory, these coauthors were unaware of Darsee's manipulations because they had little or no direct contact with the work being reported. In the case of some of the abstracts, coauthors were not even aware that the experiments reported under their names had actually been performed; sometimes they did not even know that their names were on the abstract until after it had been published.

At Harvard, Darsee's chicanery compromised the integrity of nine published papers, necessitating the withdrawal of each, in part or in toto. In addition, 21 abstracts had to be withdrawn. Darsee's supervisors in these various projects were more closely involved with the research and knew more about what was happening in the laboratory than was apparently the case in many of the projects at Emory; yet that did not prevent Darsee from violating his trust as he had done before. Nearly two years passed before he was caught cheating at Harvard, and many more months of laborious investigation were needed to obtain proof of his additional mischief.

The *Journal* published two articles by Darsee and coauthors,[2,3] both of which represent work done at Emory. In this issue, we now publish retractions of both those articles, signed by the coauthors and the dean of Emory.[4,5] The retractions are based on findings uncovered by an ad hoc review committee at Emory, and of course they were not cleared with Dr. Darsee. We would have published them in any case, but we sent copies of the retractions to Dr. Darsee to inform him of our intention to publish them and determine whether he contested the retractions. His response, also published in this issue,[6] makes it clear that he does not.

This sad affair has much to teach us all, as we consider what went wrong and whether such tragedies can be prevented in the future. One conclusion stands out: The system in academic medicine may have given Darsee the opportunity to fabricate his data, but he clearly had an inner need to do so that was quite independent of the system in which he worked. He had cheated in research as early as his college days, and he evidently continued to do so throughout his career. The explanation for his aberrant behavior surely lies to a considerable degree within his own character, and it does an injustice to those victimized by his wrongdoing to suggest that they were in some way responsible for it. I believe that all those in charge of research laboratories in academic institutions clearly have a responsibility to supervise young investigators and to set standards for their behavior as scientists. Unfortunately, faithful discharge of those responsibilities will neither protect against the occasional appearance of an unscrupulous person in the laboratory nor guarantee that his or her misdeeds will be promptly recognized. That does not mean that standards are worthless. The setting of standards is in fact crucial. Without them morale suffers and aspiring young investigators often lose their way.

Another clear message is that a clever cheat is difficult to expose. There is a widespread opinion that science is self-correcting and that fraud will be exposed relatively promptly, as other scientists attempt to repeat the original work. That view was suggested in the recent report of an Association of American Medical Colleges Ad Hoc Committee on the Maintenance of High Ethical Standards in the Conduct of Research: "The principal deterrent in research fraud," declared the committee, "is the overwhelming probability that fraudulent data will be detected soon after their presentation."[7] As a member of that committee, I now feel that this statement is too sanguine. Fraudulent data may be rapidly identified in an area of great importance where research activity is intense, but that is probably not true in most fields. I don't have all the facts, but I am not aware that many, if any, of Darsee's misdeeds were identified by the failure of others to confirm his work after its publication. Of the two retracted papers published in the *Journal*, the findings in the taurine study have not yet been challenged by other workers as far as I know. The conclusions of the study of HLA types in hypertrophic cardiomyopathy were challenged in a letter from Switzerland that we published about a year later.[8] The authors of that letter suggested various explanations for their inability to confirm the Darsee study, but they did not mention the possibility of fraud.

What kind of protection against fraud does peer review offer? The Darsee affair gives a clear answer: Little or none. Most of Darsee's fraudulent work was published in peer-reviewed journals, some with very exacting standards, and yet in none of the reviews was there enough suspicion to warrant rejection. In the case of the two papers we published, no suggestion of dishonesty was raised by any of the referees or editors, although I recall thinking at the time that some of the data seemed almost too neat. I did nothing about following up that suspicion, however, because to have challenged an author's integrity on the basis of intuition alone would hardly have been justified. A request from an editor for primary data to support the honesty of an author's findings in a manuscript under review would probably poison the air and make civil discourse between authors and editors even more difficult than it is now. The lesson seems evident: Unless a maladroit cheat fabricates results that are manifestly impossible or inherently contradictory, even the most rigorous peer review is not likely to uncover fraud.

If fraud is not easily preventable and is difficult to detect, can nothing be done about it? This question takes on special urgency when we consider that to an increasing extent collaboration among laboratories and the greater use of large teams of investigators make the prevention and detection of fraud more difficult than ever before. All investigators whose clinical,

administrative, or teaching responsibilities keep them away from their laboratories a good deal of the time (that probably includes most senior academicians in clinical departments) need to give careful thought to strengthening their working relations with their research fellows and encouraging the highest standards of intellectual honesty. Suggestions to this end are offered in several recent committee reports on the subject of fraud in science. I particularly like the sensible advice on "Prevention of Dishonesty in Science" offered by the Ad Hoc Advisory Committee to the Dean of the Harvard Medical School on Dishonesty in Scientific Research in its report of January 25, 1982. Among other things, this committee recommends "written, detailed, explicit procedures for data gathering, storage and analysis." It also suggests that "fellows should be supervised by experienced scientists in the laboratory, including regular in-depth scrutiny of the primary data and the calculations. . . ."

One final lesson from the Darsee affair is that we will all have to reconsider the meaning and responsibilities of coauthorship. The two formal retraction notices, as well as Darsee's supporting letter in our Correspondence section, seem to suggest that his coauthors at Emory had no responsibility at all for what happened, simply because they are honest and had no hand in the manipulation of the data. I cannot agree, and neither will most other editors. Anyone who allows his or her name to appear among the authors of a paper assumes major responsibilities.[9] Even if coauthors have not actually done any of the laboratory work, they should at least know that the experiments and measurements were carried out as described, and they ought to understand what was done and why. Coauthors should be able and willing to defend the paper in public, and that means they must be confident about the integrity of the data. Furthermore, coauthorship should never be conferred or accepted as an honor or simply as a reward for providing resources or sponsoring a junior investigator who has done all the work. Coauthorship should denote at least that there has been meaningful participation in the planning, design, and interpretation of the experiments and in the writing of the paper. Multiauthored papers should never be submitted for publication without the concurrence of all authors. Before accepting a multiauthored paper for publication, editors would be well-advised to assure themselves that these conditions have been met, perhaps by having all coauthors sign an appropriate form.

These considerations may help to reduce fraud, but they will never prevent it entirely. Scientists are human and therefore some of them — we hope very few — will cheat. There are ways to diminish the numbers of dishonest people who are allowed into medical research, and to minimize the damage they might do, but we need to recognize that stringent procedures designed to detect or prevent wrongdoing would probably be counterproductive. The damage done to morale and the free exchange of ideas may in the long run be far more costly than even the depredations of an occasional Darsee. In science, as in other human activities, trust has its risks, but they are far exceeded by the benefits.

ARNOLD S. RELMAN, M.D.

REFERENCES

1. Culliton BJ. Coping with fraud: the Darsee case. Science 1983; 220:31-5.
2. Darsee JR, Heymsfield SB, Nutter DO. Hypertrophic cardiomyopathy and human leukocyte antigen linkage: differentiation of two forms of hypertrophic cardiomyopathy. N Engl J Med 1979; 300:877-82.
3. Darsee JR, Heymsfield SB. Decreased myocardial taurine levels and hypertaurinuria in a kindred with mitral-valve prolapse and congestive cardiomyopathy. N Engl J Med 1981; 304:129-35.
4. Nutter DO, Heymsfield SB, Glenn JF. *Retraction of* Darsee JR, Heymsfield SB, Nutter DO. Hypertrophic cardiomyopathy and human leukocyte antigen linkage: differentiation of two forms of hypertrophic cardiomyopathy. N Engl J Med 1983; 308:1400.
5. Heymsfield SB, Glenn JF. *Retraction of* Darsee JR, Heysmfield SB. Decreased myocardial taurine levels and hypertaurinuria in a kindred with mitral-valve prolapse and congestive cardiomyopathy. N Engl J Med 1983; 308:1400.
6. Darsee J. A retraction of two papers on cardiomyopathy. N Engl J Med 1983; 308:1419.
7. Report of the Association of American Medical College Ad Hoc Committee on the Maintenance of High Ethical Standards in the Conduct of Research. J Med Educ 1982; 57:895-902.
8. Bloch A, Crittin J, Barras C, Jeannet M. Hypertrophic cardiomyopathy and HLA. N Engl J Med 1980; 302:1033.
9. Huth EJ. Authorship from the reader's side. Ann Intern Med 1982; 97:613-4.

Deeper problems for Darsee: Emory probe

The promising research career of 34-year-old John R. Darsee, MD, was aborted in 1981 when he was caught fabricating research results in the Harvard Medical School Laboratory of Eugene Braunwald, MD, chief of medicine at both Brigham and Women's and Beth Israel hospitals in Boston (*JAMA* [MEDICAL NEWS] 1983;249:1797-1807). A federal investigation done in collaboration with Braunwald and released this spring revealed that Darsee's deception contaminated five Harvard research projects, invalidating nine published reports.

Now a separate, year-long inquiry at Emory University School of Medicine in Atlanta, where Darsee excelled during his 1974-1979 postgraduate training, has uncovered a far more extensive pattern of fraud.

The scandal has touched some of Emory's most senior faculty members and administrators, including chief of medicine J. Willis Hurst, MD, who was Lyndon B. Johnson's cardiologist and a former president of the American Heart Association; Donald O. Nutter, MD, another cardiologist who is executive associate dean for research; and Robert C. Schlant, MD, chief of the school's cardiology division.

Outgoing Emory medical school dean James F. Glenn, MD, who recently became president of Mount Sinai Medical Center in New York, disclosed in an interview that he has recommended to Emory University president James T. Laney that "the faculty members involved in this be censured and that certain steps be taken to relieve them of any administrative responsibilities regarding research activities to prevent a recurrence of this sort of activity." He declined to name the faculty members he had in mind except to say that it goes beyond those whose names appear on Darsee's papers as coauthors. Glenn added that a commission should be established "to examine the cardiology program to be sure it's sound."

On May 11, at Laney's request, George W. Brumley, MD, interim dean, named a five-person committee to, in Laney's words, "frame recommendations as to what our institutional response should be, taking into account Dr. Glenn's recommendations. We will be accountable."

The examination of Emory's cardiology training program is still pending, but Hurst has asked Charles M. Huguley, Jr, MD, a senior member of his department who served on the investigating panel, to appoint a committee to "build checks and double-checks" against future research fraud.

(Glenn and others at Emory say his resignation as dean after three years was unconnected to the Darsee affair. Glenn says he stayed on longer than he otherwise might have "to finish this revolting business.")

Given the elapse of time and the absence of any primary data to evaluate, Emory's evidence against Darsee is entirely circumstantial. Nonetheless, the investigators consider it overwhelming.

Darsee's misconduct at Emory, the four-man investigative panel concluded, invalidated virtually all of the scientific abstracts that he published while in Atlanta; a majority of his full-length published papers; parts of two coauthored chapters in the 1981 edition of *The Heart*, the widely used textbook edited by Hurst, his former mentor; and an about-to-be-published cardiology manual for house staff. (The panel included Neil C. Moran, MD, Department of Pharmacology, chairman; Elmer C. Hall, PhD, Department of Biometry; Panagiotis N. Symbas, MD, Department of Surgery, and Huguley.)

Darsee, who is finishing a second residency in critical care medicine at the Ellis Hospital in Schenectady, NY, was unavailable for comment on the Emory report. He had not been provided with an advance copy of the report or with a press statement summarizing it. He was not asked to testify before the investigating panel, nor did he submit a defense or rebuttal of the allegations against him. He did send copies of what he said were original laboratory notebooks to Atlanta on request, but the investigators considered them of doubtful authenticity.

In the interim, Ellis Hospital, which had staunchly supported Darsee this spring in the wake of the NIH report of his Harvard fraud, has reconsidered his suitability as a clinician. Hospital spokesman Andrew Foster recently said that officials there felt it would be "more prudent" to remove Darsee from patient care responsibilities and restrict him to "computer applications of medical data."

"His position is under review," Foster said. "What it boils down to is that we are trying to decide, or help him decide, what his place in medicine should be, if any. Our president, William E. Schirmer, feels strongly that we have an institutional responsibility to society. If we were to say 'Goodbye, Dr. Darsee,' it becomes another community's problem, another hospital's problem."

Braunwald, who was sorely stung by the federal investigators' consensus that lax "supervisory practices" in his laboratory contributed to Darsee's misconduct at Harvard, clearly felt vindicated by the new report. "I am saddened but not surprised," he told *JAMA* MEDICAL NEWS. "It shows that his pattern of misconduct was established long before he got to Harvard."

According to the Emory investigators, the tainted research papers range from "probably fabricated" to unverifiable and presumptively invalid. One 1981 paper on myocardial taurine levels that appeared in *The New England Journal of Medicine* and was coauthored by Steven B. Heymsfield, MD, director of the Division of Nutrition at Emory University Hospital, was called by the investigators "the most flagrant example of the fabrication and dissembling that

continued on page 2871

continued from page 2867

pervades most of the papers and abstracts."

For his part, Heymsfield says he does not remember any of the individuals in the 43-member family that formed the basis of the study and believes that none were admitted to the Emory hospital's clinical research center, of which he is assistant director. Copies of documents provided to the investigating panel by Darsee contain no patient names or dates, nor do they identify the hospital or clinic from which the subjects were referred.

The study claims that HLA types were obtained for the kindred, but there are no records of these determinations. The panel found it "highly unlikely" that Darsee could have obtained samples of venous blood from all 43 persons "on two consecutive days after an overnight fast twice a year," as the paper claimed, and considered it even more implausible that he could have collected a complete set of 24-hour urine samples from all 43—including a 2-year-old child—twice a year or more frequently.

Moreover, Darsee claimed to have analyzed myocardial tissue obtained within four hours of death from three of the subjects and, by implication, from 22 controls—a feat that an Emory pathologist testified was highly unlikely. In other cases, the paper reports cardiac biopsy samples obtained during pacemaker insertion, although no one can verify that the procedures were done, and, as in the case of many of Darsee's reported clinical studies, no human studies committee approvals were applied for or obtained.

Finally, the "taurine" paper acknowledges the assistance of two physicians and a PhD biochemist who the investigators believe never existed. Another Darsee publication, from 1979, lists a coauthor, T. L. Feng, whom the Emory committee also believes to be fictitious.

Heymsfield, in an interview with *JAMA* MEDICAL NEWS, said he played a minimal role in the study and remembers protesting to Darsee that he did not deserve to be a coauthor. "Darsee said, 'You've helped me so much in the past that I insist,'" Heymsfield recalled ruefully. "Little did I realize what a trap I was getting into." Now he describes the taurine paper as "horrendous" and some aspects of the data as "really wild" but, at the time, he said, "it never dawned on me that John wasn't really doing this work. He would come into my office and talk about these patients as if he were in contact with them and say how lucky he was to have found them. I personally believe there was such a family," Heymsfield added, "but it would be incredible if he really did all he said."

An example of a fraudulent abstract is one in which 19 patients with failing myopathic hearts were allegedly taken off medication, anaesthetized for invasive tests of cardiac function, medicated experimentally, and, by implication, reanaesthetized two weeks later.

"This is a clear example of . . . pure fabrication," said Moran in an interview with *JAMA* MEDICAL NEWS. "First, it would be unlikely that any human investigations committee would approve such a study. Aside from that, it would be next to impossible for a single person to do such a complex study in a hospital without someone knowing what he's doing."

This abstract, entitled "Relationship between response to inotropic agents and homeometric autoregulation of the myopathic heart" (*Clin Research* 1980;28:163A), lists Heymsfield and Nutter as coauthors, as well as a contemporary of Darsee's, Candace L. Miklozek, MD, who is now practicing at the Beth Israel Hospital in Boston. As is the case with dozens of other Darsee abstracts, however, the Emory coauthors now say that they knew nothing about the work or the abstract at the time. Miklozek was unavailable for comment, but Emory officials say she has joined in the retraction of this and other questionable Darsee abstracts bearing her name.

A press statement issued by Emory tiptoes around the possibility that Darsee may have exposed patients to unacceptable and unmonitored risks. "None of the research reported by Darsee in full-length articles, whether actually performed or not, involved the use of drugs for human patients or the giving or withholding of therapy," the statement said, ignoring the questionable human studies reported in abstracts. Moran says he was all but convinced that these clinical studies were never actually carried out, but he acknowledges that no proof exists one way or the other.

The considerable body of work in question appeared in many of the nation's leading medical journals and was presented at the premiere cardiology meetings. Cardiologist coauthors included not only Hurst, Nutter, and Schlant, but also Nanette K. Wenger, MD, director of cardiac clinics at Emory-affiliated Grady Memorial Hospital, and Paul F. Walter, MD, professor of cardiology at Emory University Hospital.

These and other experienced researchers were so taken in by the talented, hardworking, and persuasive Darsee that virtually none of them raised a serious question about the enormous quantity of his work or its authenticity, even when abstracts began appearing that bore their names and described studies of which they had no knowledge.

Nutter, for instance, was on a leave of absence as a Robert Wood Johnson Foundation Fellow in Washington, DC, during 1978-79. Darsee would call him occasionally and send copies of charts and tables, and Nutter would give advice. When he returned to Atlanta in 1979, Nutter says he became aware that his name was on published abstracts "that I had no involvement in." On two occasions, Nutter said, he asked Darsee (who by then had gone to Harvard) to stop the practice. Darsee responded much as he had to Heymsfield's protestation. "It was a very flattering kind of thing," said Nutter. But "purely and simply, he was trading on my name."

Schlant, who has overall responsibility for the cardiology training program, told *JAMA* MEDICAL NEWS that he remembers feeling "occasional twinges"

continued on next page

continued from previous page
of suspicion about the sheer volume of research that Darsee was cranking out. But he did not raise any questions "because I could have been accused of stifling a brilliant young man," Schlant said. "He was such a popular figure and was supported so strongly by his closest co-workers. In retrospect, of course, we could have asked to see the primary data, even if it caused a lot of embarrassment to senior faculty."

Only one Emory faculty member ever articulated a doubt about Darsee's work. In 1979, when Darsee was chief resident in medicine at Grady Memorial Hospital, Ayton Someren, PhD, of the hospital's Department of Pathology and Laboratory Medicine, was disturbed by discrepancies in an abstract on the cardiac effects of pancreatitis, coauthored by Darsee and another Emory resident. Asked by the junior resident to consult on a study to be presented at Emory's annual House Officers Research Day, Someron had reviewed slides from about 20 pancreatitis cases. When she saw the resulting abstract, however, the number of cases had ballooned to 86 and the conclusion sharply differed from her own interpretation. She informed Hurst of the problem. Hurst wrote to the junior resident, with a copy to Darsee, asking them to consult with Someren about her complaint, but the matter apparently rested there.

In another case—a later abstract on septal hyper-

The paper chase

The trail begins innocently enough—just a typeset line in the 1973 Author Index of the *Cumulated Index Medicus*: Darsee JR. EKG changes and anaphylaxis [letter] South Med J 70(10):1249, Oct 77.

For the next month the index shows another letter by the same author in a different journal; two months later there is an another, and a few months later another. Rapidly the citations in the *Index Medicus* change. Letters to the editor are replaced by short, single-authored papers. Soon, longer manuscripts appear in the line listings, with two, then three, then four and five authors.

With time, the names of the journals become more prestigious, small papers give way to substantial works, and researchers of national and international stature are listed as coauthors with increasing frequency.

For the past five years, Darsee's name appears on an astounding 116 articles listed in the *Science Citation Index* (which does not include book chapters, monographs, portions of proceedings, or abstracts not published in journals). From 1978 to 1982, some 46 investigators shared authorship with Darsee on published papers. Three of these coauthors have the now dubious distinction of being listed on 28, 37, and 39 papers, respectively. Still, more than 90% of Darsee's papers list him as first or sole author; 28 of the papers list only one coauthor.

Of the 45 abstracts representing Darsee's work at Emory, 43 have now been labeled invalid. Data for the two that are thought to reflect valid research were collected and analyzed by other investigators. Of the ten full-length articles reviewed by the Emory committee investigating Darsee's work, only two are believed to represent valid research. Harvard University has retracted nine articles, in whole or in part, published by Darsee while at that institution.

In some of the publications or abstracts, absence of the original data was the primary reason for considering the work invalid. In others, the lack of records on a substantial number of cases and/or, for example, the total absence of any records on controls, suggested to the Emory committee that "they are completely fraudulent." In one article, allegedly completed at Harvard after Darsee left Emory, the Harvard work has been labeled "fictitious" by the Emory panel.

Coauthors of one article labeled invalid believe the paper is sound, but because none of them has copies of any of the original data, Emory has acted conservatively and declared the results suspect. (The coauthors are no longer at Emory.)

As for what portion of the rest of Darsee's manuscripts will be found invalid, this remains unclear, but the soundness of his papers is not the only issue in this web of scientific deceit:

The list of papers that cite Darsee's work covers more than 29 feet of computer printout from the SCISEARCH data base. The ultimate validity of these 241 papers (including 13 in which Darsee cites his own previous publications) that have cited Darsee's work as a foundation for all or part of their investigations is now open to question.

All of the coauthors of the suspect articles who still remain at Emory have, together with James F. Glenn, MD, outgoing dean of the School of Medicine, cosigned letters of retraction recently sent to the editors of the journals in which those papers appeared. In response to a specific recommendation of Emory's external review committee, each letter also is accompanied by a cover letter from the dean asking the editors to assist in disseminating this information as widely as possible.

Michael Aycock, assistant dean of the School of Medicine, says that at least one journal is looking into mechanisms for listing the retraction in its index. How or whether the articles will be identified as retracted in scientific citation systems is uncertain.

Arthur W. Hafner, PhD, Director of the Division of Library and Archival Services of the American Medical Association, says that although noting the status of the articles in an on-line system might be theoretically possible, "it is virtually impossible to modify the existing printed indexes. In many places in the world, all they have are lists of abstracts."

Eugene Garfield, PhD, president of the Institute for Scientific Information, also says that annotations for

trophy in hypertensive patients coauthored by Darsee and a different resident—Someren was listed as a coauthor, although she had not participated at all. Again she complained to Hurst, with no apparent result. Subsequently, she learned that Darsee had submitted the same material to the *American Journal of Cardiology*, with 91 instead of 78 cases.

Someren later spotted gross mislabeling of three pictures in Update IV of Hurst's text, *The Heart* (McGraw-Hill Book Co, 1981), in a chapter on neoplasia and the heart coauthored by Darsee.

In retrospect, it seems clear that the considerable damage the Darsee affair has caused would have been lessened if Someren's suspicions had been taken more seriously. But, at the time, Darsee's remarkable productivity dazzled everyone, just as it subsequently did for nearly two years at Harvard.

The Emory investigative panel was not asked to pursue the questions of how such gross deception could slip past everyone for so long and how to prevent it. But these questions were addressed in a report by a panel of three outside evaluators: David M. Brown, MD, of the University of Minnesota; Gottlieb C. Friesinger, MD, of Vanderbilt Medical Center in Nashville; and Arno G. Motulsky, MD, of the University of Washington, Seattle. Their recommendations:

continued on page 2876

---

retracted articles theoretically could be put into on-line systems but doesn't expect it to happen soon. "It could happen eventually but not easily," he says.

The eight articles that Emory University Woodruff Medical Center is retracting are:

1. Darsee JR, Heymsfield SB. Decreased myocardial taurine levels and hypertaurinuria in a kindred with mitral-valve prolapse and congestive cardiomyopathy. *N Engl J Med* 1981;304:129-135.

2. Darsee JR, Mikolich JR, Walter PF, Schlant RC. Paradoxical rise in left-ventricular filling pressure in the dog during positive end-expiratory pressure ventilation—a reversed Bernheim effect. *Circ Res* 1981; 49:1017-1028.

3. Darsee JR, Nutter DO, Hopkins LC, Heymsfield SB. Neurogenic skeletal myopathy in patients with primary cardiomyopathy. *Circulation* 1979;59:492-497.

4. Darsee JR, Mikolich JR, Nicoloff NB, Lesser LE. Prevalence of mitral-valve prolapse in presumably healthy young men. *Circulation* 1979;59:619-622.

5. Darsee JR, Heymsfield SB, Nutter DO. Hypertrophic cardiomyopathy and human-leukocyte antigen linkage—differentiation of two forms of hypertrophic cardiomyopathy. *N Eng J Med* 1979;300:877-882.

6. Darsee JR, Miklozek CL, Heymsfield SB, Hopkins LC, Wenger NK. Mitral-valve prolapse and ophthalmoplegia—a progressive, cardioneurologic syndrome. *Ann Int Med* 1980;92:735-741.

7. Darsee JR, Mikolich JR, Walter PF, Schlant RC. Trans-cutaneous method of measuring Doppler cardiac-output. 1. Comparison of trans-cutaneous and juxta-aortic Doppler velocity signals with catheter and cuff electromagnetic flowmeter measurements in closed and open chest dogs. *Am J Cardiol* 1980; 46:607-612.

8. Darsee JR, Walter PF, Nutter DO, Clapham DE. Trans-cutaneous Doppler method of measuring cardiac-output. 2. Non-invasive measurement by trans-cutaneous Doppler aortic blood velocity integration and M-mode echocardiography. *Am J Cardiol* 1980; 46:613-618.

The nine papers retracted in whole or in part by the Cardiac Research Laboratory at Harvard Medical School are as follows:

1. Darsee JR, Kloner RA, Braunwald E. Early recovery of regional performance in salvaged ischemic myocardium following coronary artery occlusion in the dog. *J Clin Invest* 1981;68:225-239.

2. Kloner RA, DeBoer LWV, Darsee JR, Ingwall JS, Braunwald E. Recovery from prolonged abnormalities of canine myocardium salvaged from ischemic necrosis by coronary reperfusion. *Proc Nat Acad Sci* 1981;78:7152-7156 (retracted in part).

3. Braunwald E, DeBoer LWV, Glogar DH, Darsee JR, Ertl G, Ingwall JS, Kloner RA. Observations on the experimental reduction of infarct size and on the delayed consequences of transient myocardial ischemia. *Acta Med Scand* 1981;210/suppl 651:123-132 (retracted in part).

4. Darsee JR, Kloner RA, Braunwald E. Time course of regional function after coronary occlusions of 1 to 120 minute duration. *Am J Physiol* 1981; 240:399-407.

5. Kloner RA, DeBoer LWV, Darsee JR, Ingwall JS, Hale S, Tumas J, Braunwald E. Prolonged abnormalities of myocardium salvaged by reperfusion. *Am J Physiol* 1981;241:591-599 (retracted in part).

6. Darsee JR, Kloner RA, Braunwald E. Demonstration of lateral and epicardial border zone salvage by Flurbiprofen using an in vivo method for assessing myocardium at risk. *Circulation* 1981;63:29-35.

7. Kloner RA, Darsee JR, DeBoer LWV. Reducing infarct size with coronary reperfusion. *J Cardiovascular Medicine* 1981;67:1173-1181 (retracted in part).

8. Darsee JR, Kloner RA. Dependency of location of salvageable myocardium on type of intervention. *Am J Cardiol* 1981;48:702-710.

9. Darsee JR, Kloner RA. No reflow phenomenon: a time-limiting factor for reperfusion after coronary occlusion? *Am J Cardiol* 1980;46:800-806.

Each letter of retraction from Harvard contained a detailed account of the specific portions of the paper that were retracted but also listed the conclusions that were confirmed by repeat experiments in the Cardiac Research Laboratory.

—by Bruce Dan, MD, 1983 Morris Fishbein Fellow

continued from page 2873

- Trainees should be supervised "in an individualized fashion" by "identifiable mentors," and "no trainee, however exceptional, should be excused from ongoing supervision by one [or where appropriate, more than one] experienced faculty member."
- Rigorous scrutiny of approvals for human studies should be carried out before the studies begin.
- All reports and abstracts should be reviewed before submission, and ongoing research should be reviewed on a regular basis "regardless of the purported success or failure of the project. Such reviews must include inspection of the original data."
- Research fellowship programs, especially large ones like Emory's, "must be designed in such a way that the intense and frequent personal interchanges among all participants are not compromised."
- Research records should be accessible to investigators "before and after publication of the results."

Enlarging on this, Brown noted that the geographically and institutionally fragmented nature of Emory's cardiology training program—with residents scattered among six different sites—hampers "close supervision and cohesive communication among faculty members in reference to complex research protocols and techniques." As explained by Michael Aycock, assistant dean, Darsee was able to deceive Emory for so long because "he would tell Preceptor A that Preceptor B had approved something or seen the data, and then tell Preceptor C that A and B had approved it."

Beyond these recommendations, individual researchers at Emory are pondering what might be done to salvage their wounded reputations and tighten the systems of accountability that failed so spectacularly in Darsee's case. Opinions are not uniform, but, says Moran, "the real lesson for science is that everybody on a research team has to be completely involved intellectually—not to catch fraud per se but to assure high quality science. If that ideal were observed, I don't see how a Darsee could succeed at all."—by RICHARD A. KNOX

APPENDIX 6

OVERVIEWS OF MISCONDUCT AND FRAUD IN RESEARCH

Misconduct in Research: An Issue of Science Policy and Practice

DARYL E. CHUBIN

SCIENTIFIC activity is now carried on as a profession. Like all professions, it requires the mastery of a large body of knowledge and it certifies practitioners to engage in that pursuit by the award of degrees and by academic appointments. Like other professions, it has a code of conduct, although the ethical code of the scientist has not been formally and explicitly promulgated by any official organisation of the profession. As in all professions, breaking the code that governs professional conduct is regarded with severe disapproval. Allegations of cheating, dishonesty and misconduct in research carry a social stigma. The stigma may be retroactive and contagious; it creates the apprehension that dishonesty in research by a few will contaminate the reputation of the whole body of scientific knowledge and the whole profession of science.

The misconduct of a few scientists threatens to discredit the profession in the eyes of the public, if the public believes that misconduct in research has been discovered and has then been ignored. At a time when scientists and scientific research are already being criticised by persons who identify science with technology and who deplore some of the consequences of technology, dishonesty among scientists causes unease among scientists themselves and regretful or gleeful misgivings among publicists who are critical of science.

Misconduct in Research: Causes and Consequences

Exact data on the scale of misconduct in science are difficult to obtain. Suspicions have often gone unexplored and local investigations into misconduct have not been reported. In all professions there is a tendency—perhaps to a fault—to grant the benefit of the doubt. The trust that binds members of the scientific community, and especially those who labour in the same laboratory, also tends to throw a protective veil over those who are suspected of misconduct.

While the ethos of science is based on arrangements such as control by peers and professional autonomy, the ethic of liberal democracy starts from a different set of assumptions. "Instead of peer control and autonomy, there is public control and accountability; instead of internalised standards, there is public scrutiny; instead of self-regulation and self-evaluation, there are checks and balances, external regulations, and publicly produced evaluations."[1] This difference is significant in an age of expensive,

[1] Prewitt, Kenneth, "The Public and Science Policy", *Science, Technology, & Human Values*, VII (Spring 1982), pp. 5-14.

governmentally financed "big science." Dr Kenneth Prewitt, president of the Social Science Research Council in the United States, concludes that:

> It is too late to call for a "separation of science and the state." It is too late for science to evade the consequence of participatory democratic culture in which it is practised and from which it draws support. . . . The contract is not being renegotiated because an attentive public is mobilised against science. It is being renegotiated because science is of public consequence.[2]

Autonomy and accountability are not mutually exclusive. They are in a state of continuous compromise. The cases of misconduct in research that emerged in the 1970s precipitated a vigorous discussion. Interpretations have ranged from generalised arguments for the view that the cases known in public are only the "tip of the iceberg", to the view that the discovered incidents of misconduct are the inconsequential and mischievous work of a "few rotten apples". Publicists who write about science tend towards the former interpretation; scientists towards the latter.

If we examine the cases of those accused in recent years of misconduct in research—Summerlin, Straus, Long, Alsabti, Soman, Spector, Darsee—and the institutions where they occurred—the Sloan-Kettering Institute and Boston, Yale, Cornell, Harvard and Emory universities—the dim outline of a pattern emerges. The majority of these episodes occurred in biomedical research and at eminent institutions. There has been reluctance to investigate the charges, or to make them public. Why? Some would answer that such misconduct is simply the work of a few demented minds. This may be a partial explanation, but it is certainly not sufficient to explain the phenomenon. In order to understand—but not to excuse—the misconduct of scientists as scientists, it may be more useful to reflect on the structure of contemporary scientific practice.[3] We do not consider in our analysis certain kinds of misconduct, such as fabrication of data under pressure of a desire for monetary profit in trials of clinical drugs, purloining of data and hypotheses from documents submitted in connection with applications for financial support of research, and claiming the credit in publications for research done by others.[4]

In addition to the ambiguity of professional and ethical norms and the inadequacy of the transmission of these norms during graduate training,[5] at least seven explanations of misconduct in scientific research have been

[2] *Ibid.*, p. 13.

[3] Broad, William J., "Fraud and the Structure of Science", *Science*, CCXII (10 April, 1981), pp. 137–141; Chubin, Daryl E., "They Blinded Us with 'Science'?", *Science, Technology, & Human Values*, XIII (Fall 1983), pp. 23–29.

[4] E.g., Holden, Constance, "FDA Tells Senators of Doctors Who Fake Data in Clinical Drug Trials", *Science*, CCX (1980), pp. 749–750; Marshall, Eliot, "Federal Courts Finds IBT Officials Guilty of Fraud", *Science*, CCXXII (4 November, 1983), p. 488.

[5] Cowen, Robert, "The Not-So-Hallowed Halls of Science", *Technology Review*, LXXXVI (May-June 1983), pp. 8, 85.

offered: psychopathy,[6] unbridled ambition,[7] pressure for publication,[8] competition for federal support,[9] the "lab-chief" system,[10] failings of the "peer-review system",[11] and lack of replication and sheer sloppiness.[12]

This enumeration focuses not only on individual malefactors but also on the various institutions which—through appointment, publication, the allocation of resources and the bestowal of honours—vouch for the integrity of scientific work. The institution in which research is done confers a presumption of probity. This tends to forestall, through incredulity and inertia, the examination of "deviant" episodes which occur within them, such as the misdeeds of Darsee at Harvard.[13] Scientific institutions tend to treat each detected misdeed as an isolated incident with no general implications; this happened in the case of the rampant plagiarism of Alsabti.[14]

Few efforts have been made to estimate, on the basis of reports by the scientists of privileged knowledge, the extent of misconduct in laboratories. One survey conducted in Great Britain obtained about 200 reports of "intentional bias", including the "purposeful manipulation of an experimental design or results to confirm an hypothesis".[15] Although the population from which this sample was drawn is unknown, two findings are instructive. One concerns the amount of "bias" detected by different methods. It shows that one third of the cases were detected through the presence of suspected data; a quarter of the detections stemmed from difficulties of replication; 3 per cent of those reporting claimed they were never detected; 45 per cent of the cases of misconduct involved more than

[6] Jones, F.L., "Obsession plus Pseudo-science equals Fraud: Sir Cyril Burt, Intelligence, and Social Mobilty", *Australian and New Zealand Journal of Sociology*, XVI (March 1980), pp. 48–55; McDonald, Kim, "Fraud in Scientific Research: Is It the Work of Psychopaths?", *The Chronicle of Higher Education*, XXVI (15 June, 1983), p. 7.

[7] Broad, William J., "Would-be Academician Pirates Papers", *Science*, CCVIII (27 June, 1980), pp. 1,438–1,440, and "Imbroglio at Yale (I): Emergence of a Fraud", *Science*, CCX (3 October, 1980), pp. 38–41; Wade, Nicholas, "Madness in their Method", *The New Republic*, CLXXXVIII (27 June, 1983), pp. 13–17.

[8] Jackson, C. Ian and Prados, John W., "Honor in Science", *American Scientist*, LXXI (September–October 1983), pp. 462–464.

[9] Hixson, Joseph, *The Patchwork Mouse* (Garden Cty: Doubleday, 1976); West, Susan, "Crackdown on Lab Fraud", *Science 83*, IV (May 1983), p. 14.

[10] Broad, William J. and Wade, Nicholas, *Betrayers of the Truth: Fraud and Deceit in the Halls of Science* (New York: Simon & Schuster, 1982).

[11] Relman, A.S., "Lessons from the Darsee Affair", *New England Journal of Medicine*, CCCVIII (9 June 1983), pp. 1,415–1,417.

[12] Hawkes, Nigel, "Tracing Burt's Descent to Scientific Fraud", *Science*, CCV (17 August, 1979), pp. 673–675; Schmaus, Warren, "Fraud and Sloppiness in Science", *IIT Perspectives*, I (September–December 1981), pp. 1–4.

[13] Butterfield, Fox, "Harvard Suspends Doctor for Fraud", *The New York Times*, 16 December, 1981; "Harvard Disputed on Fakery", *ibid.*, 25 January, 1982; "Report Absolves Harvard Faculty of Blame in Fakery by Researcher", *ibid.*, 26 January, 1982. Culliton, Barbara J., "Fraud Inquiry Spreads Blame", *Science*, CCXX (25 February, 1983), p. 937.

[14] Broad, W. J., "Would-be Academician Pirates Papers", *op. cit.*; and "Jordanian Denies He Pirated Papers", *Science*, CCIX (11 July, 1980), pp. 38–41.

[15] St James-Roberts, Ian, "Cheating in Science", *New Scientist*, LXXII (25 November, 1976), pp. 98–101.

one investigator, 15 per cent having at least three collaborators, and some of those having ten or more collaborators. Thus, the act of deception which is thought to be the work of a single individual may often, upon closer scrutiny, be an act of collusion, or even—where the misconduct was planned in advance—of conspiracy.

The United States National Institutes of Health (NIH) reported, between 1980 and 1982, 45 cases of alleged misconduct which had come to their attention. Viewed against a total of more than 20,000 grants and contracts each year, the fraction is miniscule.[16] By the summer of 1983, an NIH institutional liason officer was receiving in the office of extramural research "at least a couple [of calls] each month".[17] These usually came from alarmed co-workers with first-hand knowledge of a situation; an average of one call per month warranted further investigation.[18]

One estimate of the extent of a variety of forms of misconduct in research stems from a recent analysis content of the published papers written in co-authorship by the cardiologist John Darsee when he was working at the Harvard and the Emory University medical schools.[19] While this analysis unearthed some striking findings in a non-random sample, and is hardly an indictment of biomedical research practices, it suggests that cheating in the form of "involuntary co-authorship", "re-publication of abstracts", with minimal changes in title and text, and "lack of candour" in recording research protocols, are widespread parts of "acceptable practice". At the very least, this study supports the thesis that errors and sloppiness were rife in the 109 collaborative papers published between 1978 and 1982 and bearing Darsee's name. In many papers, the deficiencies were blatant, yet reviews by Darsee's co-workers and the referees of scientific journals either failed to detect them or did not regard them as sufficiently serious to prevent publication.

These efforts to estimate the magnitude of the various patterns of misconduct in research are unsatisfactory. The data neither span the whole range of the sciences nor are they systematically collected; they centre on case histories in bio-medical research. Thus, we must depend on the accounts of journalists and the co-operation of technicians, co-authors, programme-officers and editors to fill in the blanks between accusation and the admission of wrongdoing. Few blanks have been filled in, since colleagues have wished to keep matters "in-house", whether within the institutions, the discipline, the journal or the profession.

[16] Brandt, Edward T., Jr. "PHS Perspectives on Misconduct in Science", *Public Health Reports*, XCVIII (March–April 1983), pp. 136–139.

[17] Colt, George Howe, "Too Good to be True", *Harvard Magazine* (July–August 1983), pp. 22–28.

[18] Miers, Mary, personal communication, May 1984. For various views of suspected misconduct and procedures for its investigation, see Swazey, Judith P. and Scher, Stephen R., (eds), *Whistleblowing in Biomedical Research: Policies and Procedures for Responding to Reports of Misconduct* (Washington, DC: USGPO, 1982).

[19] Stewart, Walter W. and Feder, Ned, "An Examination of Certain Professional Practices of 47 Biomedical Scientists", unpublished manuscript.

Practice and Malpractice in Research

To explain the causes of misconduct in research requires an examination at each of the four stages through which a scientific investigation must pass. The four stages are: production, reporting, dissemination and evaluation. Misconduct can occur at each stage.

At each stage, normal activities predominate. At each stage, and especially at the first two, there are opportunities for misconduct, for deviations from the prevailing code which is as comprehensive as the entire range of the activities which make up research, but which is nevertheless general and vague.[20] Deviations from the code must be intentional; incompetence is not the same as misconduct. Nevertheless, it is difficult to discern intentions in a wholly reliable way.[21]

Research is intended to investigate new ideas or the rediscovery of old ones which had been exhausted. Originality is therefore crucial in the initiation and conduct of research. Given the high appreciation of originality in science, those who appropriate another scientist's words or propositions are stealing the one thing that will come into common possession if the idea turns out successfully, but which must be credited to one or more particular persons or teams who may legitimately claim have originated it. The offence of theft is repugnant and it seldom occurs in a crude form.[22] Yet the origin of a scientific idea is not always easy to ascertain.[23] The attribution of discovery to particular individuals for inspiring or developing scientific ideas which are regarded as original contributions becomes more difficult as teamwork in the laboratory creates collective claims to share credit for the discovery of particular ideas. The purloining of ideas or data is certainly a deviation from the ethical code of science, as is claiming credit for a discovery made by someone else. Yet theft of data or misappropriation of credit are not the forms of misconduct which concern me here. I am interested in the misconduct which occurs in the falsification of data. This means reporting falsely about observations and methods and claiming to have obtained results by falsely reported observations, about observations which have in fact not been made.

[20] Schmaus, Warren, "Fraud and the Norms of Science", *Science, Technology, & Human Values*, VIII (Fall 1983), pp. 12–22; Zuckerman, Harriet, "Norms and Deviant Behaviour in Science", *Science, Technology, & Human Values*, IX (Winter 1984), pp. 7–13.

[21] E.g., Racker, Efraim, "The Warburg Effect: Two Years Later", *Science*, CCXXII (21 October, 1983), p. 232. On reconstructing intentionality, see Zuckerman, Harriet, "Deviant Behaviour and Social Control in Science", in Sagarin, Edward (ed.), *Deviance and Social Change* (Beverly Hills: Sage, 1977), pp. 87–138; Weinstein, Deena, "Fraud in Science", *Social Science Quarterly*, LIX (March 1979), pp. 639–652.

[22] Broad, J., "Would-be Academician Pirates Papers", *op. cit.*; "Imbroglio at Yale (I)", *op. cit.*; "Jordanian Denies Pirated Papers", *op. cit.*; "Imbroglio at Yale (II): A Top Job Lost", *Science*, CCIX (10 October, 1980), pp. 171–173; "Must Plagiarism Thrive?", *British Medical Journal*, CCLXXXI (5 July, 1980), pp. 41–42.

[23] E.g., Nelkin, Dorothy, "Intellectual Property: The Control of Scientifc Information", *Science*, CCXIV (14 May, 1982), pp. 704–708.

Teamwork operates in the collection and analysis of data. There are procedures for making measurements and for making entries into data books, but there are few hard-and-fast rules. Within the division of labour carried out in laboratories, the "honour system" is the preferred mode of operation.[24] The practice of "cross-checking" at the bench varies from institution to institution and from team to team; expediency and convenience often determine the norm; there are "local" traditions in such matters. In their haste to produce another paper, to demonstrate a technique, or to support an hypothesis, scientists sometimes verge on deviant practices.[25]

The fabrication of data, either partially, as in the case of Cyril Burt[26] or with wholesale "forgery"—"dry-labbing"—as in the case of Darsee,[27] are bolder and more desperate acts than the "trimming" of data to favour a certain outcome; this procedure was observed by Charles Babbage over a century ago and there have been occasional occurrences ever since. Scientific heroes such as the physicist Millikan[28] and charlatans alike[29] have "massaged" raw results into confirmations of their preferred theories. Such enhancements may indeed be seen as integral to the creative process.

To say that Galileo, Mendel and Newton all "fudged" their data is not to indict the science of their day or to tarnish their reputations. Rather, it reminds us that "prevailing standards" of procedure, evidence and proof change to recalibrate the threshold of the "finagle factor".[30] Revisionist history fails to adjust expectations to practice;[31] such a portrayal of scientific standards does violence to the epoch under scrutiny.[32]

[24] Knorr-Cetina, Karin D., *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science* (Oxford: Pergamon, 1981), ch. 4–5.

[25] This inference is made, e.g., in the report on Darsee's misdeeds in Braunwald's Harvard Medical School laboratory at Brigham and Women's Hospital. See, *Report of the National Heart, Lung, and Blood Institute Special Panel to Review Alleged Misconduct at Brigham and Women's Hospital/Harvard Medical School, Boston, Massachusetts*, 29 June, 1982.

[26] Dorfman, D. D., "The Cyril Burt Question: New Findings", *Science*, CCI (29 September 1978), pp. 1,177–1,186; Gillie, Oliver, "Burt's Missing Ladies", *Science*, CCIV (8 June, 1979), pp. 1,035–1,039.

[27] Culliton, Barbara J., "Emory Reports on Darsee's Fraud", *Science*, CCXX (27 May, 1983), p. 937.

[28] Holton, Gerald, "Subelectrons, Presuppositions and the Millikan–Ehrenhaft Dispute", *Historical Studies in the Physical Sciences*, IX (1978), pp. 161–224; Franklin, Allan, "Millikan's Published and Unpublished Data on Oil Drops", *Historical Studies in the Physical Sciences*, XI (1981), pp. 185–201.

[29] See Gould, Stephen Jay, "Morton's Ranking of Races by Cranial Capacity", *Science*, CCXX (5 May, 1978), pp. 503–509.

[30] Gould, Stephen Jay, "The Finagle Factor", *Human Nature*, I (1978), pp. 80–87.

[31] Wade, Nicholas, "Scandal in the Heavens: Renowned Astronomer Accused of Fraud", *Science*, CXCVII (18 November, 1977), pp. 707–709; Dolby, R.G.A., "Reflections on Deviant Science", in Wallis, Roy (ed.), *On the Margins of Science: The Social Construction of Rejected Knowledge* (Keele: University of Keele, 1979), pp. 9–49.

[32] Joravsky, David, "Unholy Science", *The New York Review of Books*, XXX, (13 October, 1983), pp. [illegible]5; Ziman, John, "Fudging the Facts", *The Times Literary Supplement*, No. 4,197 (9 September, 1983), p. 955.

Contemporary practice features countless versions of the scientific method. Generalisations about scientific methods can be misleading.[33] A not uncommon phenomenon is "self-deception". The history of science is littered with examples of how strong beliefs—*pro* or *contra*—can alter perceptions. "Premature ideas", such as the theory of continental drift, that are eventually confirmed, experimental observations that long mislead scientists before they are proved illusory, for example, polywater N-rays, and claims that evoke such endless debate that they become virtually synonymous with "pseudoscience", such as the salutary effects of vitamin C and the existence of *psi*-phenomena—all bear traces of self-deception.

The word "hoax" is often applied to episodes of self-deception and fraud in research. What is often overlooked in discussing the hoax, however, is that the knowingly false claims have been sustained by the beliefs of reputable scientists. From the manipulation at the bench to interpretation of statistical analyses at some distance from the bench, scientists are capable of seeing and writing what they want to believe is happening.[34] Work in a laboratory does not provide immunity from such self-deception; it may do just the opposite by promoting undue resistance to new ideas or, alternatively, a fervour that supports dubious claims. Such "deceptions" wax and wane with the accepted beliefs of a part of the scientific community; it is on this collectivity that the burden of proof lies.[35] Professor Stephen Jay Gould contends that scientists are ignorant of their own behaviour: "The problem is that most scientists have bought into the mythology; they think that they're pursuing something objectively and they're not sufficiently self-critical."[36]

Cyril Burt's fabrication of data on IQ is a good case in point. Dorfman's painstaking re-analysis of Burt's data shows the construction of those data "from a theoretical normal curve, from a genetic regression equation, and from figures published more than 30 years before Burt completed his surveys".[37] Burt's eminence, plus an unrelenting style in disarming critics, no doubt forestalled the discovery that the increasing number of twins reported in his writings after 1950 were a product of his imagination, his desire to defend "true belief", and paranoia: "His secrecy, suspicion of rivals, egocentricity, compulsive motivation, and hypochondria are all consistent with this diagnosis. But whether Burt actually deluded himself, and believed his own fictions, is harder to answer."[38]

Perhaps the most compelling evidence comes from the journalist who first used the word "fraud" in reference to Burt's work. Mr Oliver Gillie suggests

[33] Collins, Harry, "The Seven Sexes: A Study in the Sociology of a Phenomenon, or the Replication of Experiments in Physics", *Sociology*, IX (May 1975), pp. 205–224.

[34] Rosenthal, Robert and Rubin, Donald B., "Interpersonal Expectancy Effects: The First 345 Studies", *The Behavioral and Brain Sciences*, I (September 1978), pp. 377–415.

[35] Pinch, Trevor J., "Normal Explanations of the Paranormal: The Demarcation Problem and Fraud in Parapsychology", *Social Studies of Science*, IX (August 1979), pp. 329–348.

[36] Canadian Broadcasting Corporation, "Science and Deception: Part II", transcript of "Ideas" broadcast, 17 October, 1982, p. 6.

[37] Dorfman, D. D., *op. cit.*, p. 1,177.

[38] Hawkes, N., *op. cit.*, p. 675.

that to lend additional credibility to his research, Burt invented co-workers. He used the names of real persons and attributed work to them that they did not do. As a postscript to the exposé of Burt by Kamin in 1974,[39] Mr Gillie writes, "Burt's whole corpus of work must now be suspect".[40] He then goes on to ask why Burt's work was looked at so uncritically by psychologists and others—notably, Eysenck, Jensen, Darlington, Jinks, Roberts and Shockley—for so long. "The answer might tell us something important about the role that power, charisma and wishful thinking can play in bolstering support for scientific theory."[41] Scientists are indulgent—at least for a time—of their "great men".[42]

In reporting of their own research results, scientists are conscious of the conventions for selecting and presenting information. These conventions are rooted in disciplinary and laboratory traditions and in the individual's training and style of research. Drafts for presentation to readers beyond their most immediate circle of colleagues must reduce a corpus of thought and data to acceptable specifications of length, tone and significance. The presentation in laboratory seminars comprises the first test that these specifications have been met.

There are gaps, however, between the expectations of immediate leagues and those of the larger scientific community. The biased summary of arguments in the literature, the "cooking" of data to favour hypothesis *a* over hypothesis *b*, and the omission of relevant sources in the literature, are all acts of misrepresentation. To the outsider, they appear to be deception; to the author—an individual or a team—they may be innocent errors. To distinguish error from deception is difficult but necessary.[43] No senior author, editor or reviewer wants to entertain the possibility of "cooking", or "apparent lack of candor".[44] In the case of Philip Felig and Vijay Soman, supervision was lax. Soman was a postdoctoral fellow who plagiarised data and text from the unpublished manuscript of a scientist working under a grant from the National Institutes of Health. This information was contained in a manuscript sent to Felig to review for publication. The manuscript was not published, but the data ended up in a paper which Soman had drafted with Felig as an allegedly involuntary co-author; he claimed to have had no knowledge of the plagiarism. The National Institutes of Health reluctantly called for an inquiry. As a result, Soman was forced to leave scientific work, Felig lost his appointment at a

[39] Kamin, Leon, *The Science and Politics of IQ* (Potomac, MD: Erlbaum, 1974).

[40] Gillie, O., *op. cit.*, p. 1,038.

[41] *Ibid.*, p. 1,039.

[42] For another British case study of a hoax perpetrated by an accomplished, eccentric physicist and tolerated by his colleagues, see Wynne, Brian, "Between Orthodoxy and Oblivion: The Normalization of Deviance in Science", in Wallis, R., *op. cit.*, pp. 67–84.

[43] Broad, W. J., "Would-be Academician Pirates Papers", *op. cit.*, and "Imbroglion at Yale (I)", *op. cit.*

[44] Stewart W. W. and Feder, N. *op. cit.*

famous medical school, and Yale University suffered the indignity of a formal inquiry and a censure.[45]

Not all improprieties are as clearly established as that of Soman, and, presumably, of Felig. This was a simple breach of the prohibition of plagiarism. It did not involve the falsification of reports of observations; it simply claimed the credit for observations made by others. The false description of the techniques used in research is another matter, and closer to the heart of the matter—likewise, the false recording of observations and their false reporting. Here it is not always easy to distinguish fabrication from sloppiness or negligence. Two recent cases, one Swiss, the other American, illustrate these problems; the first of these cases, is especially relevant to the issue of the determination of fraud in research.

Dr Karl Illmensee, a Swiss development biologist, was credited with the distinguished achievement of transplanting cell nuclei between the embryos of mice in 1980. "Irregularities in the recording of experimental protocols", in research at the laboratory of the University of Geneva, were charged in 1983 by a team at the Wistar Institute in Philadelphia.[46] The latter team was unable to replicate the procedure used by Dr Illmensee nor could his collaborator at Bar Harbor, Dr Peter Hoppe. An international commission found no evidence of systematic fraud, but numerous "errors and discrepancies". The suspicion of manipulation of experimental protocols cast a shadow over Dr Illmensee's pending application for a three-year renewal of his grant from the National Institutes of Health: "Loss of financial support could make it difficult for Illmensee to carry out the commission's recommendation that he repeat the challenged experiments."[47] Without funds, Dr Illmensee could not repeat the experiments and the National Institutes of Health thereupon declined to renew the grant, even though no evidence of fraud had been found. It did so on the grounds that there were enough problems "to throw grave doubts on the scientific validity of the conclusions".[48] Dr Illmensee's reputation has been damaged, and consequent difficulties in obtaining financial support for his research may prevent him from re-establishing his once high reputation.

The second case—one of presumed plagiarism—is that of Professor Kenneth Melmon, the chairman of the department of medicine at Stanford University. He is charged with incorporating "large chunks of copyright material from a book he helped to edit into a chapter he wrote for another textbook. The material was used without attribution and apparently without

[45] Hunt, Morton, "A Fraud that Shook the World of Science", *The New York Times Magazine*, 1 November, 1981, pp. 42–46.

[46] Marx, Jean L., "Bar Harbor Investigation Reveals No Fraud", *Science*, CCXX (17 June, 1983), p. 1,254; "'Not Guilty' Says Biologist Charged with Fraud", *New Scientist*, XCIX (28 July, 1983), p. 252.

[47] Norman, Colin, "Illmensee Faces Funding Cutoff", *Science*, CCXXIV (20 April, 1984), p. 265.

[48] Holden, Constance, "NIH Turns Down Illmensee Proposal", *Science*, CCXXV (6 July, 1984), p. 36.

permission".[49] A committee on ethics of the Stanford University medical school found Professor Melmon guilty of "grossly negligent scholarship," i.e. of negligence, and not of "deliberate plagiarism". He was censured for "carelessness in dealing with the intellectual property of others"; he resigned his chairmanship, but retained his professorship, so as not to "weaken or destroy the work that I had done over the past five years".[50]

There is reluctance to charge misconduct in recording and reporting in the explanation of unaccountable research results. Error is assumed, and therefore tolerated as the explanation, yet it is at the stage of recording and reporting that really serious misconduct occurs. Failures in the selection and presentation of data and the description of method are violations of the conventions of research which editors and reviewers sometimes detect and expose. But as Dr Arnold Relman, editor of the *New England Journal of Medicine*, has said: "You can't run science the way you run a criminal investigation."[51] Reporting may or may not be a precise mirror of practices in the laboratory.[52]

The susceptibility of research to fraudulent action is characteristic of the dissemination of the results of research. It can occur at any of the three semi-public steps, and in publication in the scientific literature. The first step—the formal release of new results to scientific peers through drafts, papers ("preprints"), telephone conversations, correspondence and presentations at colloquia and seminars in "friendly" laboratories—has become a feature of work in "big science". The "peers" to whom these communications are made may be former teachers, students and co-workers, as well as respected members of the scientific community, including the influential scientists who head competing laboratories. They can be trusted to respond in a timely and candidly critical fashion to the new information. These responses are intermediate evaluations beyond the immediate limits of the particular laboratory in which the work is done, prior to the submission of the paper for publication. They are the first and perhaps first line of defence of the integrity of scientific knowledge from fraudulent deformation. Yet they are not always effective.

The stage of publication is an additional line of defence of the integrity of scientific knowledge, but it is also a line at which infringements on the ethics of scientific conduct can occur. Editors and referees have traditionally had the responsibility of assessing the scientific soundness of the papers submitted to them. The first task of an editor or a referee is to detect assertions of observations and interpretations which are false, incomplete or logically unpersuasive. But this responsibility carries with it the opportunity

[49] Norman, Colin, "Stanford Investigates Plagiarism Charge", *Science*, CCXXIV (6 April, 1984), p. 35.

[50] Norman, Colin, "Melmon Resigns Stanford Chairmanship", *Science*, CCXXIV (23 June, 1984), p. 1,324.

[51] Relman, A. S., *op. cit.*

[52] Faberge, A. C., "Open Information and Secrecy in Research", *Perspectives in Biology and Medicine*, XXV (Winter 1982), pp. 263–278.

for misconduct. Advisers and members of panels of peers who review applications for the support of research have similar responsibilities, and they also have opportunities for the same kind of misconduct, namely, blocking the progress of scientific ideas and of scientists who are distasteful to them and of purloining their ideas and data.

Although a submitted paper and an application for support are considered confidential documents, the ideas and data they contain can be "pirated" or stolen and then published under the name of the "pirate". A strong recommendation not to publish the paper, or to require far-reaching revision as a condition for publication, or to reject a proposal for support, permits a lapse of time during which a referee or adviser can utilise the acquired knowledge for his own benefit. The anonymity of referees and peers fosters the possibility of abuse. The burden thus falls largely on editors. As Dr Thomas Stossel, editor of the *Journal of Clinical Investigation*, asks: "When suspicions are aroused, how much raw data can they demand as a condition for publication without alienating prospective authors?"[53] Repeated incidents may be the only compelling evidence that something untoward has occurred.

The misappropriation of ideas or results is more likely to be done by the reviewers of research proposals than by editors and referees. A survey of 300 cancer research workers provided some evidence of piracy by reviewers of data contained in grant proposals submitted to the National Institutes of Health that were refused support; 20 per cent of those who submitted proposals claimed that their data had been misappropriated by reviewers.[54]

Referees and peers are also competitors. Some cheating in competition for scarce federal funds allocated by "peer review" must, by now, be regarded as not wholly improbable. The dishonesty and formulation of ways to deter and detect such misconduct is a large challenge.

In "peer review" in journals and on behalf of grant-awarding bodies, the identity of the author or the institutional affiliation serves to skew evaluation; this happens even in so-called "blind reviews".[55] There is a favourable presumption about the honesty or quality of manuscripts submitted from the laboratories of the more eminent institutions; it is a variant of the "Matthew effect".[56]

Can we call a reviewer or adviser negligent when he defers to the reputation of the author or of his institution? What about an author who fabricates bibliographical references or is unwilling to supply editors with raw data or "missing" references? To raise such questions in the course of

[53] Stossel, Thomas, personal communication, May 1984.

[54] Gillespie, Gilbert W., Jr, Chubin, Daryl E. and Kurzon, George, M., "Experience with NIH Peer Review: Researchers' Cynicism and Desire for Change", *Science, Technology & Human Values*, X (Summer 1985).

[55] Harnad, Steven (ed.), *Peer Commentary on Peer Review: A Case Study in Scientific Quality Control* (Cambridge: Cambridge University Press, 1982).

[56] Merton, Robert K., "The Matthew Effect in Science", *Science*, CLIX (5 January, 1968), pp. 56–63.

the process of review would introduce the very distrust that editors insist would impede the creativity of research. On the other hand, the reluctance of authors to be forthcoming and of reviewers to be vigilant engenders delay and suppression and makes all the parties involved culpable to some extent.[57] Acts of misconduct are not confined to editors and referees. Applicant authors are also susceptible to the opportunities for misconduct at this stage.

One such act of misconduct lies in the attribution of co-authorship. Unacknowledged as well as falsely attributed collaboration can be injurious. If authorship is considered a status earned through some contribution to the finished product, and not an honour bestowed for "remarks made in an elevator"—as one irate "co-author" has observed—then the criteria for allocating credit and responsibility for scientific substance of a paper should be clear-cut. In the rules governing recording and reporting, the rules governing the attribution of authorship should be clear-cut. They are, however, not so clear-cut, as editors and others know. Nor can any uniform criteria be imposed by journals since the traditions of particular institutions militate against such joint action.[58]

Another infringement on the ethics of scientific activity on the part of the authors is the simultaneous submission of the same paper for publication to more than one journal. There is certainly some deceit in such behaviour, since each editor is permitted by the author to construe the submission as assigning the right of first refusal to his journal. An author engaging in such a practice may claim it is to safeguard against unwarranted delay in reaching an editorial decision and eventual publication; protracted delay in the acceptance of manuscripts for publication is the bane of authors in branches of science which are moving rapidly. Yet the question is whether the pursuit of priority of publication which, after all, defines "originality", justifies the practice, or even an isolated instance, of multiple submission.

When a paper is finally published, the *imprimatur* of various institutions is given to it. These include the author's own institution and the unit within which he has worked, the agency or foundation which supplied financial support for the research, the journal and its editors and referees, who represent a sponsoring professional association or society or a commercial publisher. All the responsibilities, risks and benefits associated with authorship or co-authorship rest on the published paper.

A failure to acknowledge participants in the conduct and completion of a piece of research—from laboratory technicians to anonymous referees advocating revisions prior to publication—also raises ethical questions, but it is difficult to draw clear lines of obligation. The opportunities for abuse

[57] Kilbourne, Brock K. and Kilbourne, Maria T. (eds), *The Dark Side of Science*, Vol. 1, Part 2 (San Francisco: AAAS, Pacific Division, August 1983); Culliton, B. J., "Emory Reports on Darsee's Fraud", *op. cit.*

[58] For example, see "ACS Guidelines for Publishing Research Proposed", *C&E News* (26 September, 1983), pp. 39–43.

increase as the piece of research progresses. As the number of persons who participate in the research and its dissemination increases, the precise lines of accountability become blurred.[59]

Misconduct redounds to the discredit of the institutions in which it occurs just as distinguished achievements redound to their credit. These institutions share responsibility for the quantity and integrity of the research done or published under their auspices, as they do for the confidentiality and non-abuse of knowledge in the process of its discovery and publication. When charges of misconduct in research are made, the "collective responsibility" to protect, investigate, correct, compensate or punish, must be specified.

Once published, a paper is supposed to contain the essential information that permits replication by another research worker. Confirmation or refutation is then expected to follow, and thus—so it is believed—the scientific community carries forward its work. But this is not easy. The fragmentation of knowledge that has accompanied the emergence of new specialities increases the necessity for scrutiny of new results, but it also renders the scrutiny by the larger scientific community more difficult.[60]

Replication is performed far less routinely than textbooks suggest. Because replication is unoriginal, it is also unrewarded. A diversion of scarce resources is unlikely unless some previously staked intellectual territory is in jeopardy of "subdivision" or "take-over" by scientists possessing new tools or data, or unless a discovery is imminent. As one research worker in cancer has stated: "in a field in which there are many scientists interested . . . one soon finds out what the real answers are."[61] Replication is a kind of defence mechanism. When "important" scientific problems preoccupy teams far removed in space from each other, the scientific community recognises the timeliness of getting "real answers".[62] As the primary literature of any scientific specialty attests, an extension of knowledge must be sought, not its mere duplication. Collective evaluation will determine what replication is necessary, by whom, and in what medium.

The silent suspicion of "peers" regarding the inadequacy, inappropriateness or errors they detect in the methods, findings and interpretations of published papers are left undocumented. Why do the "peers" not speak out? Are they too busy? Do they fear being wrong? One conjecture is that they do not wish to injure the self-esteem or otherwise damage those they judge by alleging that they have been incompetent. The

[59] Croll, Roger P., "The Noncontributing Author: An Issue of Credit and Responsibility", *Perspectives in Biology and Medicine*, XXVII (Spring 1984), pp. 401–407.

[60] Weinstein, Deena, "On the Importance of Fraud in Science", *IIT Perspectives*, I (September–December, 1981), pp. 4–6; Chubin, Daryl E., "Commentary (on the Schmaus–Weinstein Exchange)", *IIT Perspectives*, II (March–June 1982), pp. 11–12.

[61] Canadian Broadcasting Corporation, *op. cit.*, Part I, p. 12.

[62] E.g., see Kolata, Gina Bari, "Reevaluation of Cancer Data Eagerly Awaited", *Science*, CCXIV (16 October, 1981), pp. 316–318.

benefit of such vigilance to the vigilant is small. There are a few powerful persons and many fragile reputations in science.

Misconduct in Science before the Lay Public

Scientific knowledge is not confined to the scientific community. It is also before a wider audience. What constitutes misconduct in disclosure before an uninformed public? Accounts in the popular press are especially subject to the magnification of "abuses" by excessive simplification and misinterpretation. Such misrepresentation, especially in the daily press, is not necessarily malicious—it is a consequence of the journalists' need to inform the public and to sell the news.

"Fraud in science" has been a popular topic, especially in *The New York Times*, the national newspaper of the United States which gives the largest amount of attention to science.[63] Sometimes it is well reported. Reporting by local newspapers is sometimes excellent, as were the various short articles published in the *Atlanta Constitution and Journal* on Darsee's misconduct at Emory University.[64] However, if we look to Sunday supplements and magazines, quality is quickly diluted. These stories are not necessarily the handiwork of sensational journalists either, although the accounts in the *Washington Post* of the cold reception to Broad and Wade's *Betrayers of the Truth* by the scientific community are examples of hyperbole.[65]

Although science is thought to get no better or worse treatment in quality of coverage than any other major social activity,[66] what distinguishes science-writing for the lay public is the frequency of articles, their topical focus and the selection of the sources used.[67] The very images that headlines conjure up can lead to unfounded beliefs that "science", like "government", is the home of felons. If the images contained in a few column inches are negative, then they cease to be the newsworthy exceptions; they seem to damn scientists at every turn. But this means that scientists incur new obligations to inform reporters rather than dismiss or

[63] Tumbleston, Robert, "Scientific Misconduct and its Portrayal in Newspapers", unpublished manuscript, Cornell University, May 1984. On national coverage of the Darsee case, see Butterfield, Fox, "Doctor Ousted by Harvard is also Suspected of Falsifying Research at Emory", *The New York Times*, 24 April, 1983; "Emory Tells How Physician Faked Much Research", *ibid.*, 7 May, 1983; "Research Drew Suspicion at Emory", *ibid.*, 19 May, 1983.

[64] Seabrook, Charles, "Doctor's Research at Emory in Doubt", *Atlanta Constitution*, 4 April, 1983; "Emory Jolted by Faked Heart Research", *ibid.*, 6 May, 1983; "Emory Probe Includes Top Heart Experts", *ibid.*, 17 May, 1983; "Emory Urged to Clamp Down on Researchers", 27 May, 1983; "Heart Research Fraud Cuts at Emory's Reputation", *Atlanta Journal and Constitution*, 29 May, 1983; "Publication Retracts Two Emory Studies", *Atlanta Constitution*, 9 June, 1983; "Scientific Deceit", *Atlanta Weekly*, XXXIV (19 June, 1983), p. 18.

[65] Slung, Michele, "A Book Some Scientists Love to Hate", *Washington Post Book World*, 20 February, 1983, p. 15; see also Wade, N., *op. cit.*

[66] Goodfield, June, *Perspectives on Science and the Media* (Washington: AAAS, 1981).

[67] Ryan, Michael, "Attitudes of Scientists and Journalists toward Media Coverage of Science News", *Journalism Quarterly*, LVI (Spring 1979), pp. 18–26, 53; Goodell, Rae, "Problems with the Press: Who's Responsible?", *BioScience*, XXXV (March 1985), pp. 151–157.

exploit them. If they do not co-operate, then they cannot complain so bitterly about misquotations, intrusions and the persistent lack of public understanding of science in general and of their own research in particular. Making one's work intelligible is a vital part of accountability. Learning to do this, however, is not taught in graduate schools.

Responses to Misconduct in Research: The Apportionment of Responsibility

Responses to misconduct in research range from the annoyance and embarrassment of certain scientists to investigations by special panels of "impartial administrators", to the expulsion of the dishonest scientist and the requirement of restitution of the misused funds.[68]

Sometimes the discovery of misconduct leads to new prescriptions for remedial action, but these are largely unenforceable in practice. It is true that recognition of the misconduct can lead to public and private stock-taking, and this may in turn lead to altering practices by increasing the moral self-discipline of scientists. Arousing and forming public opinion of the scientific community is a prerequisite for inhibiting, detecting and punishing.

It is said that misconduct retards the progress of scientific knowledge, for example, through wasting the time of those who have to investigate the charges, and through the cancellation of what was for a time thought to be a genuine addition to knowledge. Retraction of published articles undermines the mutual trust required in research. The making of misleading claims to knowledge causes subsequent research on their clinical applications to be wasted effort. Misconduct in research is also a waste of public money and it increases demands for fiscal accountability. It leads to demands for more stringent scrutiny by external, non-scientific institutions, especially the legislature and its subsidiary bodies and by the media of mass communication.

Misconduct in relatively secondary matters such as plagiarism, false claims to authorship, the misappropriation of data and ideas from proposals submitted for financial support or papers for publication, are also damaging to the standing of science and to the morale of the scientific community.

Misconduct in research raises questions about more rigorous rules for detecting it, responding to it when detected, and preventing its recurrence. It also raises questions about the adequacy of current laboratory arrangements for the formation and training of young research workers.

Responsibility for the actions which should follow in disclosures of misconduct is, one might argue, that of scientists alone. But the onus of "collective responsibility" falls also on the administrators of the various

[68] Woolf, Patricia, "Fraud in Science: How Much, How Serious?" *Hastings Center Report*, XI (October 1981), pp. 9–14.

institutions implicated through their sponsorship of research in which misconduct is alleged. These institutions include universities and independent and affiliated research institutes, the scientific journals, professional scientific societies, the governmental bodies which pay for most research, the legislative body which allocates the governmental research budget, the media of mass communication which disseminate the results of research, and the courts which adjudicate on disputes over the negligence and liability for misdeeds. All these bodies, in differing ways, see themselves as responsible for punishing the malefactors and preventing the recurrence of the malefactions for which they may be held accountable. University officials have been shocked by and indignant over the discovery or confession of irregularities in research—at their own institutions and elsewhere. The unfolding of events in the case of Dr Darsee, which is the most publicised and expensive malefaction to date, shows the inadequacy of procedures at a major institution of research and teaching, the Brigham and Women's Hospital, which is affiliated with Harvard University.

Dr Darsee was a promising cardiologist in the laboratory of Professor Eugene Braunwald, who called him "brilliant" and "creative". Dr Darsee admitted to fabricating data that tainted the research done with nearly three quarters of a million dollars of governmental funds. Over 50 papers and abstracts were retracted. As the evidence was uncovered, it led back to the medical school of Emory University where more falsification of data by Dr Darsee came to light. Professor Braunwald's response to the furore over handling of the case, including a six-month review[69] by a "blue-ribbon committee" which issued the so-called Ross report, was: "I got a bum rap."[70] His excuse for the sluggish internal investigation was: "If I had gone to NIH and publicly exposed him, it would have ended his life. If he can't apply for a grant, he's a dead man."[71] Professor Braunwald benevolently granted him the benefit of the doubt:

> In the belief that Darsee had committed a single foolish act, Braunwald and other Harvard officials elected not to notify others when Darsee was first caught in the spring of 1981. NIH senior staff who reviewed the panel's report [in December of that year] wrote that the "desire to be fair to Dr Darsee and to ensure due process is commendable, but . . . a large and costly study of great importance for a major public health problem was irrevocably compromised because of the failure to inform" [NIH and co-investigators].[72]

By the time the review panel of the National Institutes of Health released its report, the issue had become not the testimony of Dr Darsee and his

[69] "Report of the Ad Hoc Advisory Committee to the Dean of the Harvard Medical School on Dishonesty in Scientific Research", (Ross report), *Harvard University Gazette*, LXXVII (29 January, 1982), pp. 11–12; Mosher, Ellsworth H., "Fraud Investigation", *Science*, CCXV (19 March, 1982), p. 1,456.

[70] Culliton, Barbara J., "Coping with Fraud: The Darsee Case", *Science*, CCXX (1 April, 1983), pp. 31–35.

[71] Colt, G. H., *op. cit.*, p. 24.

[72] Culliton, B. J., "Emory Reports on Darsee's Fraud", *op. cit.*, p. 937.

co-workers, but "the relevant practices, procedures, and supervision that existed" in the laboratories of the Harvard University medical school.[73] Whereas the Ross report had exonerated all but Darsee, the panel found that Harvard, Dean Tosteson of the medical school, Professor Braunwald and Dr Kloner—a research worker at Brigham and a co-author with Darsee—were "insufficiently rigorous" since they limited their investigation of Darsee's work to an "internal audit" which "may have contributed to the delay in discovering the extent of the data falsification". An external investigation "could have led to an earlier resolution of these unfortunate events".[74]

At Emory University, Dr Darsee was co-author with the most prominent research workers in the medical school, including its associate dean. Names were not removed from "completely fictitious" abstracts, for example, because Darsee was "fluent and convincing" that credit was due for training received.[75] None of Dr Darsee's co-authors at Emory has ever spoken publicly about his share in Darsee's fraud nor has any one of them expressed contrition over the fabrications in which he was, according to Dr Darsee, an unknowing accomplice. The full text of the investigation by the medical school at Emory University was released to the press.[76] The investigation of Emory University by the National Institutes of Health concluded that Darsee's "research activities had not been adequately supervised by senior faculty . . . and that no one person was aware of the totality of his overall research efforts".[77] Furthermore, several of Darsee's papers covering work done at Emory University had been submitted for publication after he had moved to Harvard. This was somewhat unusual for a postdoctoral fellow which is what Darsee had been.[78]

Dr Darsee's activities at Emory and Harvard Universities generated an outburst of more general suspicion. By the summer of 1983, 80 per cent of American medical schools, according to one estimate, moved to adopt rules for investigating falsifaction.[79] This estimate differs markedly from that of a Washington lawyer that less than "one per cent of all American colleges and universities, which differ categorically from medical schools, have adopted formal policies and procedures for misconduct investigations".[80] Several

[73] *Report of the National Heart, Lung, and Blood Institute Special Panel*, *op. cit.*, p. 3.
[74] *Ibid.*, p. 30.
[75] Culliton, B. J., "Emory Reports on Darsee's Fraud", *op. cit.*, p. 936.
[76] "Report of Ad Hoc Committee to Evaluate Research of Dr John R. Darsee at Emory University" (Moran report), press release, 28 March, 1983 (addendum released 5 May, 1983).
[77] "Report of the Division of Research Resources (DRR) Ad Hoc Consultant Group to Examine the Conduct of Dr John R. Darsee's Research at the Emory University School of Medicine, General Clinical Research Center (EUSM, GCRC)", NIH Memorandum, 27 March, 1984, p. 2. See also, DeBlieu, Jan, "The Darsee Affair", *Emory Magazine* (December 1983), pp. 6–15.
[78] Moran, Neil, personal communication, July 1984.
[79] West, S. *op. cit.*
[80] American Association for the Advancement of Science, "Professional Ethics Group Reviews Guidelines for Misconduct Investigations", unpublished draft, June 1984.

major universities—Brown, Massachusetts Institute of Technology, Stanford and Yale, among others—have drafted policies on "academic fraud". A stated policy may be little more than a gesture, but it reinforces the role of professional societies and federal agencies. Most of all, such actions by universities can reduce the *ad hoc* character of ensuing investigations. They are indicative of a preparedness to act more effectively and more speedily in the future.

Governmental agencies have responded actively to the disclosure of misconduct. The National Institutes of Health has drafted rules outlining procedures and punishments for contractors and recipients of grants who are responsible for irregularities in research.[81] In the case of Dr Darsee, two sanctions were imposed: his disbarment from support by the National Institutes of Health for ten years, and the close review for one year of the supervisory practices and research procedures at Professor Braunwald's cardiac research laboratory. This latter action was tantamount to formal censure. In addition, the National Institutes of Health requested, and received in 1983, the repayment of $119,000 of a project-grant to Harvard for research which had been declared useless because of invalid data. These represent the harshest penalties ever imposed by a governmental agency on a research worker and his institution for the falsification of data since the procedures for disbarment were enacted in 1980; the power to disbar has been included in executive orders since 1965.[82]

In addition, the internal investigative unit of the National Institutes of Health now has an "alert system" for tracing, through agencies of the Public Health Service, individuals and institutions who have been suspected of misconduct.[83] This system can be consulted at any time, especially prior to the granting of a new award. Individuals are informed when they are being placed in the system and can request access to their file under the Privacy Act.[84]

The "alert system" challenges scientific and academic institutions to devise ways of protecting their members who may be accused of wrongdoing while complying with the disclosures that the federal patron will demand. When, for example, should an announcement be made that an internal investigation is underway? Without a resolution of the dilemma over the "balancing of interests"—and, simultaneously, the responsibility for candour—there will be recurrences of the case of Dr Marc Straus at Boston

[81] "Public Health Service Policies for Handling Misconduct in Science: Operating Policies and Procedures for Awarding Agencies and Programs", mimeographed draft, 20 October, 1983.

[82] "Government-Wide Debarment and Suspension Procedures", Hearings before the Subcommittee on Oversight of Government Management of the Committee on Government Affairs, US Senate, 97th Congress (Washington: USGPO, 11–12 March, 1981).

[83] "Public Health Service Alert System for Institutions, Organizations and Individuals under Investigation or Subject to Post-Investigational Sanctions," mimeographed draft, 21 October, 1983.

[84] Brandt, E. T., *op. cit.*, p. 138.

University. Dr Straus claimed that there had been a conspiracy by his staff and he filed suit to clear his name. His is the only recent case where the accused admitted no wrongdoing whatsoever. Mark Spector, a student of Professor Efraim Racker at Cornell University, and Darsee have admitted some guilt but they have denied the majority of the charges against them.

The Straus case was the centrepiece of the Senate Labour and Human Resources Committee hearings in 1981 into the abuse of contracts awarded by the National Cancer Institute.[85] Dr Straus was already under investigation by the National Institutes of Health and the Food and Drug Administration, and the *Boston Globe* had run a series of articles a year earlier about his:

> million-dollar clinical empire . . . alleged to have been partly based on falsified data, which, according to several team members, had been doctored on the specific orders of Straus and also because of "general anxiety" that a shortage of statistically acceptable patients might threaten future funding for many of their clinical research programs. Because of the muddle of faked and real data, some team members feared that wrong treatments were being administered.[86]

The resignation of a physician on Dr Straus's team precipitated charges laid before administrators at Boston University regarding Straus's "complicity". He was asked to resign. Dr Straus claimed he was the victim of a conspiracy and "that evidence against him had been faked".[87] Speaking publicly for the first time about the case, at a special hearing of the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research in 1981, Straus said "he had been waiting for peer review of the charges . . . [by] specialists and principal investigators in clinical oncology, like himself".[88] This complicated case and the lawsuit are still unresolved.

Journal editors have been agitated no less than governmental agencies about allegations of misconduct in research. One editor recently summarised the lessons from the Darsee affair. He cited no practical alternative to the "presumption of honesty in research, because it would be impossible to verify every primary datum and every descriptive statement in a research report".[89] His statement implied scepticism about the common view that assessment by referees or peers is a certain protection from the publication of fraudulent data. A recent experiment with the re-submission, under assumed names, of published papers in psychology, has demonstrated that the simplest measures of quality control can be circumvented: editors and referees can be duped.[90]

[85] Sun, Marjorie, "Hatch Batters NCI with Straus Case . . .', *Science*, CCXII (19 June, 1981), pp. 1,366–1,367.

[86] Broad, William J., ". . . But Straus Defends Himself in Boston", *Science*, CCXII (19 June, 1981), pp. 1,367–1,369.

[87] *Ibid.*, p. 1,368.

[88] *Ibid.*

[89] Relman, A. S., *op. cit.*

[90] Harnad, S., *op. cit.*

"The swell of scientific articles has taxed the efficacy of the referee system."[91] This "swell" of articles may be accounted for in part by the "publication-by-the-pound" syndrome and the division of experimental results into "least publishable units". Prolific publication is a by-product of the competition for appointments which confer prestige, at least in the biomedical field. As long as publication is the "coin in the realm", the refereeing system will be taxed and journals will be deluged by papers, even though their rates of rejection of submitted manuscripts fall short of the 90 per cent which is now the rate among journals in the social sciences.

Whatever the explanation for the present proliferation of papers submitted for publication, journals cannot be "the sole checkpoint for the prevention of fraud. . . The task of 100 per cent detection is monumental, overwhelming, and probably impossible".[92] One editor has declared that "any attempt to police the system would probably create such an atmosphere of suspicion and distrust that the free spirit of scientific inquiry would be crushed".[93] Dr Benjamin Lewin, the editor of *Cell*, has echoed this article of faith:

> the same spirit of trust that prevails when a paper is written and submitted to the outside world for scrutiny tends to prevail within a laboratory. If someone in your laboratory says to you this is the experiment I did, these are the reagents I used, and this is the result I obtained, you are not likely to doubt it.[94]

Yet nothing will replace "the diligent referee with the excellent memory".[95] Editors must assume that a few deceiving authors will penetrate their system. But everyone should at least be notified of the rules, as the American Chemical Society does in its "Ethical Guidelines to Publication of Chemical Research" by outlining the obligations of journal editors, authors and reviewers. Perhaps Sigma Xi's recognition of the fallibility of scientific publishing is the most fitting of all:

> publish or perish is ubiquitous across disciplinary boundaries and has no regard for the character of the institution. It has been intensified by a drop in the birthrate, which has produced a highly competitive academic environment in which employment opportunities are created mainly by death, retirement or failure to achieve tenure. There is little prospect of a self-correcting mechanism. Few institutions have the power or inclination to resist; the individual scientist . . . usually has neither.[96]

Governmental administrative agencies and the legislative branch have demanded more vigilance in what one dean of a medical school has called the "inherent conflict between surveillance and creativity".[97] Control over

[91] Colt, G. H., *op. cit.*, p. 27.
[92] La Follette, Marcel C., "Journal Peer Review and Public Policy", *Science, Technology & Human Values*, X (Winter 1985), p. 5.
[93] Relman, A. S., *op. cit.*, p. 1,415.
[94] Canadian Broadcasting Corporation, *op. cit.*, p. 10.
[95] La Follette, M. C. *op. cit.*, p. 5.
[96] Jackson, C. I. and Prados, J. W. *op. cit.*, p. 464.
[97] West, S. *op. cit.*, p. 14.

the supply of funds for research is the major device available to those who allocate public revenues. This is why the National Institutes of Health, according to its associate director, Dr William Raub, "put the onus for prevention and detection on the institution for misusing grant money or falsifying reports".[98] The misdeeds of one research worker can jeopardise the future of the entire institution.

In an attempt to increase the severity of the penalties for misconduct in research supported by public funds, congressional committees, such as the Subcommittee on Investigations and Oversight of the Committee on Science and Technology, the House of Representatives, have censured the grant-awarding bodies for their indulgent treatment of those charged with dishonesty. In the hearings on fraud in biomedical research in the House of Representatives in 1981, the view was succinctly put: "It took ten days for Boston University to investigate Straus and demand his resignation. Yet 22 months later you NIH officials are still giving him a grant."[99] While the National Institutes of Health speak cautiously about "blacklisting", legislators regret the readiness of study sections to separate allegations of past misconduct from the merits of pending proposals for new grants.

The 98th Congress moved to make the statutes effective. Section 485 of the Health Research Extension Act of 1983[100] concerns "protection against scientific fraud". The act would require that:

> (a) each entity which applies for a grant, contract, or cooperative agreement . . . for any project or program which involves the conduct of biomedical or behavioral research submit . . . assurances . . . that it (1) has established an administrative process to review reports of scientific fraud . . . and (2) will report to the Secretary of the Department of Health and Human Services any investigation . . . which appears substantial. (b) The Director of NIH shall establish a process for the prompt and appropriate response to information provided . . . respecting scientific fraud in connection with projects for which funds have been made available under this Act.[101]

President Reagan vetoed the bill containing these provisions in 1984.

Professional associations such as the Association of American Medical Colleges, which in 1982 in its *The Maintenance of High Ethical Standards in the Conduct of Research* was the first to propose rules for investigating misconduct, and Sigma Xi and the American Chemical Society, which are the most recent to do so, have historically "taken honesty in research as something so vital that it has not needed mention".[102] Many influential associations and societies now have rules like those of the Association of American Medical Colleges (AAMC).

[98] Broad, W. J., "Fraud and the Structure of Science", *op. cit.*, p. 140.

[99] Broad, William J., "Congress Told Fraud Issue 'Exaggerated'", *Science*, CCXII (24 April, 1981), p. 421.

[100] "Health Research Extension Act of 1983 (H. R. 2350) (Title IV)", US Senate and House, 98th Congress, 1st Session (18 November, 1983). The pertinent section is Part F, Sec. 485, "Protection Against Scientific Fraud".

[101] *Ibid.*, pp. 106–107.

[102] Jackson, C. I., and Prados, J. W., *op. cit.*, p. 462.

The Association of American Medical Colleges has suggested, for example, that allegations should be investigated in confidence, first by departmental chairmen, deans and select committees, then disclosed to collaborators and sponsors of the accused through a second, more formal investigation.[103] No time is stipulated for the completion of these steps, although speed is recommended. It says that no assumption that the conduct in question is an isolated event should be made; disclosure to former as well as to present collaborators and employers is proposed. It is aware of the consequences to which investigations might give rise:

> It is recognised that in these procedures a faculty member's reputation is put at risk. . . This is justified since scientists on the university faculty occupy a special place of privilege and responsibility and must be held to a higher standard of conduct. The procedures must indeed be fair to the individuals involved. They must be designed to be responsive to the special responsibility that science and faculty have to society.[104]

The report of the Association of American Medical Colleges emphasises the joint responsibility of the institution and its staff. The difficulty in achieving a true partnership in policy and practice was recognised in the report of the committee on integrity of research of the Association of American Universities (AAU), when it said "that policies and procedures must be consistent with the institution's policies on academic governance, freedom, responsibility, and due process as well as with legal restraints".[105] Charges of misconduct against a member of the teaching staff in projects supported by federal funds expand the role of the university from that of an institution of learning to that of an agent of the government.

The media of mass communication—press, radio and television—also have responsibilities. In addition to "science-writers" in the daily press and popularisers of science in magazines like *Science 84*, there has grown up a considerable body of publicists of science; some are or were scientists—others, without being scientists, have given much attention to science policy. Some of these writers by reporting that something is amiss in science, suggest that this is proof that the ideal has never existed and that faith in the "quest for truth" has been long misplaced.[106]

This style is represented in an article on the case of Dr Darsee in the widely circulated *Newsweek*. "A Case of Fraud at Harvard" featured in bold relief: "Brazen fakery of a 'golden boy' besmirches the labs in yet another

[103] The full text of the AAMC guidelines is contained in "Report of Ad Hoc Committee on the Maintenance of High Ethical Standards in the Conduct of Research", *Journal of Medical Education*, LVII (1982), pp. 895–902.

[104] Culliton, Barbara J., "AAMC Speaks of Coping with Fraud", *Science*, CCXVII (16 July, 1982), p. 226.

[105] Association of American Universities, "Report of Committee on Integrity in Research", undated mimeograph, (1982?), p. 6.

[106] Wade, Nicholas, "A Diversion of the Quest for Truth", *Science*, CCXI (6 March, 1981), pp. 1,022–1,025.

scientific scandal."[107] The supply of such copy is apparently in demand, as is shown by the following advertisement in *Science*,[108] in the columns devoted to the announcement of available appointments and fellowships:

INQUIRY
Scientific Fraud
Major news organization looking into
instances of scientific fraud occurring
in recent past or currently.
P.O. Box 767
Ansonia Station
New York, N.Y. 10023
Let us know how to reach you. Confidentiality
guaranteed.

The courts may be the authority of last resort to resolve the question of liability for scientific misconduct. Yet to view scientific fraud as a special case of "legal fraud" defined in tort law is fruitless.[109] Proving harm or injury of a physical, economic or emotional nature, resulting from intentional negligent misrepresentation or deceit in science would be a formidable task.

Perhaps only the federal government would enter the courts to recover dishonestly used research funds. This has not been done. No academic institution would welcome the publicity such a case would bring. Yet the non-legal alternatives gain force by the imminent threat of a case in court.

Nevertheless, both the American Association of Medical Colleges and the National Institutes of Health use language in their rules which implies that the courts may become the eventual arbiter of conflicts over scientific misconduct. For example, the Association urges that academic institutions adopt policies which "define misrepresentation of research data as a major breach of contract",[110] while the National Institutes of Health, in trying to separate "ethical questions from those of technical liability" and "outright fabrication" from the "overzealous interpretation of data", has pondered the legal liability of members of study-sections on "matters of libel and slander".[111]

An argument could be made for treating misconduct in research as a crime involving theft or misappropriation of some kind.[112] The federal government is most apt to press for restitution but individuals who are falsely accused or discredited might also have recourse to the courts. If

[107] "A Case of Fraud at Harvard", *Newsweek*, XCIX (8 February, 1982).

[108] The advertisement appeared in the issue of 29 April, 1983, p. 546.

[109] Shaw, Margery Wayne, "Legal Aspects of Scientific Fraud", A. D. White Professor-at-Large Lecture, Cornell University, 26 September, 1983.

[110] Culliton, B. J., "AAMC Speaks of Coping with Fraud", *op. cit.*, p. 227.

[111] Broad, William J., "NIH Grapples with Misconduct", *Science*, CCXVII (16 July, 1982), p. 227.

[112] Kilbourne, B. K. and Kilbourne, M. T., *op. cit.*, pp. 38–43.

medical malpractice suits and "whistle-blowing" in engineering firms are omens of things to come, then scientific malpractice cases might well enter the courts in the not so distant future.

The consequences of misconduct and the apportionment of responsibility for it should not obscure its place of origin and occurrence, which is the profession of science. It is, after all, the practice of science, and not its ethical codes or threatened sanctions, that justify to scientists the defence and scrutiny of their conduct. The revision of practices is at the heart of the procedures proposed for the detection and deterrence of misconduct in research. Such new procedures are sought to close the gap between the norms which are espoused and the daily practice of research. "The very notion of being a professional implies an acceptance of moral responsibility for the consequences of one's work which affect both the other members of the profession and society at large."[113]

Curtailing and Controlling Misconduct

If what occurs in a laboratory cannot be externally regulated, then the policies, codes, rules and injunctions requiring good scientific practices which emanate from the professional associations might turn out to be no more than sounding brass or a tinkling cymbal.

Nevertheless, stringent monitoring and management alone are not likely to be sufficient. If we weigh the words of some of the commentators, monitoring can also erode both the trust that underlies the activity of scientific inquiry and the public's confidence in the truthfulness of its intellectual results. How does the scientific community deter conduct which it considers aberrant, without undue interference at the points where misconduct occurs?

Individuals and institutional administrators must discourage misconduct. This is problematic, however, because interpretations of the words "aberrant" and "undue" are made by many independent and separate institutions. They define, deter and admonish according to the custom and context of each institution. Second, these efforts constitute *ad hoc* policies the meaning and enforceability of which wane when seen outside their immediate social context. Rules do not long survive when transplanted; they become effective as mechanisms of control over conduct when they are integrated into the outlook and actions of the persons who inhabit the institution and express its expectations about the practice of research. A set of universally shared prescriptions and prohibitions is, at best, a guide to conduct. But the will to enforce prescriptions must operate in the daily life of the laboratory.

Control in the laboratory: Four procedures, which could be observed at the

[113] Goodfield, J., *op. cit.*, p. 63.

bench, have been prescribed by the National Institutes of Health to the 75 clinical centres it supports.[114] One is the use of multiple experimenters, especially the principal investigators and the heads of their laboratory. Not only is the presence of authority a deterrent, but, more positively, it has an exemplary effect. The visibility of an active senior scientist conveys to staff his interest in and awareness of the workaday details of research.[115] It also reduces the hierarchy that grows in large laboratories with numerous projects and principal investigators, many postdoctoral fellows and large auxiliary staffs. It diminishes the distance between research administrators and research workers.[116]

Accountability begins with measurements made at the bench. In biomedical research, these measurements are inscribed in data-logs or data-books and could be kept open in the laboratory and accessible to all participants—however defined—in the project. The logs should be retained for about five years after the work is completed, according to the National Institutes of Health, for inspection upon request by other scientists, editors and university administrators to support the honesty of claims to knowledge. Such a check on validity could be seen as an optional condition for publication imposed at the discretion of an editor or a university committee on the ethics of research.

Explicit rules for conduct in the laboratory could be established in written form and even displayed. Such rules would, in the words of the American Association of Universities, be the "absolute responsibility" of the director of the laboratory; and it is "inadvisable that they be delegated". They would initiate research workers into the culture and history of the laboratory and state the expectations for experimental protocols and the "assignment of credit and responsibility", including co-authorship.[117] The rules would direct members of the scientific staff to the appropriate persons for clarification when uncertainty arose.

Control in the institution: The continuity and articulation of policy and action between a laboratory or department, and the larger institution of which it is a part, is vital. Procedures to consider might be the utilisation of review committees to assess the design of research proposals, including protocols for the collection of data and the use of the time of graduate students, postdoctoral fellows and technicians; the utilisation of these measures could be expanded. Such an internal control should proceed hand-in-hand with reviews of the ethical problems of experimentation with human subjects.

[114] Norman, Colin, "Reduce Fraud in Seven Easy Steps", *Science*, CCXXIV (11 May, 1984), p. 581.

[115] Association of American Universities, *op. cit.*, p. 5.

[116] See Chubin, Daryl E., "Mentors and Students: Social Heritability in Science", unpublished ms, Georgia Institute of Technology, 1980; and Stetten, DeWitt, Jr, "Reported Laboratory Fraud Biomedical Sciences", *Science* CCXXVI (21 December, 1984), pp. 1,375–1,376.

[117] *Ibid.*, pp. 4–5.

The aim would be to identify research-designs, the execution of which presents opportunities for abuse.

The development of such procedures in scientific institutions could co-ordinate federal regulations with practices within the institutions. This might lead to the designation of an administrator to review research operations throughout the university, which would enlarge the scope of responsibilities for the vice-president for research. It would lend the authority of the university as a whole to these procedures.

Among the responsibilities of such a position would be the development of procedures for dealing with the suspicions of "whistle-blowers". Brown University and the Massachusetts Institute of Technology favour a standing committee on academic dishonesty. Such committees would have to be sensitive both to the accused and the accuser because the process of investigation could affect the morale of the entire university:

> The need to preserve confidentiality in the proceedings, respecting the rights to privacy by the individual accused of misconduct, and the duty of the institution to inform others, including colleagues, students, and funding sources [must be observed]. . . When research data and laboratory note-books are necessary evidence in reviewing the merits of the complaint, to whom do such materials belong—the individual scientists or the host institution?[118]

Such questions need to be dealt with locally, bearing in mind the requirements of the scientific community throughout the country and the external patrons. Now that scientific societies have promulgated and made public rules concerning integrity in research, scientific and academic institutions must adapt these prescriptions and to their own practices.

Control by scientific societies and journals: Professional societies should publish in their own journals or newsletters revised policies regarding the multiple submission of papers, the attribution of co-authorship, and the allocation of space for articles that derive from the same project and are either under submission or have been published elsewhere.

A more ambitious undertaking would be an examination of the practice of peer review in journals, including the rights and responsibilities of editors and referees. Regardless of the system employed, a letter to referees on the desirability of pursuing suspicions of untoward behaviour such as the manipulation of data or interpretation might be written by editors. The possibility of devious conduct on the part of the contributors must be imprinted on the consciousness of those who apply the journal's criteria for publication.

Control by governmental agencies and private foundations: Every governmental agency and private foundation that makes grants on a competitive basis could develop and publicise rules regarding misconduct in research. Three agencies of the United States Public Health Service have

[118] American Asssociation for the Advancement of Science, *op. cit.*, pp. 3–4.

already done so. Such policies might range from the administration of reprimands to individuals and the suspension of a project or its termination, to the disbarment of the individual from future consideration for grants, and the withholding or refusal of future grants to an institution unwilling to investigate alleged misconduct by one of its members. At the same time, care must be taken to deal fairly with suspected individuals. Objections have been raised to the endorsement by the Association of American Medical Colleges of the following position of the National Institutes of Health: "The university has an obligation to notify a granting agency when the responsible officer concludes that there is substantial reason to believe that fraud has occurred, even if the investigation has not been completed."[119] This statement contradicts guarantees of confidentiality. One critic has written: "If universities pursue the policy recommended in the report of the Association of American Medical Colleges, they are playing directly into the hands of those who have no conscience concerning the politicisation of the universities."[120] Dr Herman Wigodsky concludes: "One would hope that the AAMC and AAU would be acting to head off or to minimize federal regulation, not to supplement it"![121]

Governmental agencies should reassess their review of procedures for dealing with proposals for support. Existing rules should be amended to anticipate and reduce the probability of abuses by membership of study-sections or panelists selected *ad hoc*, who use their access to the privileged information contained in research proposals for their own benefit, while acting as agents or consultants of a governmental grant-awarding body.[122]

A large part of the business of the legislature is to hear testimony on issues which concern the public. Scientists participate in legislative inquiries because they seek continued support for research in general and for their research in particular. They should more willingly interpret their craft to the curious, and sometimes hostile, laymen who are guardians of the public interest. The era of denial of scientific misconduct came to an end in 1981 with the testimony of the late Philip Handler, then president of the National Academy of Sciences.

Legislative hearings are often conducted in a glare of publicity. The press can make "events". Rules for reporting cases of misconduct in a way which should give some understanding of the norms and practices of the scientific community and which do not glorify, sensationalise, mislead or persecute, would be valuable. Scientific practices conform to different rules from those

[119] Stated in Wyngaarden, James B., "Recommendation for Debarment—Decision", NIH Memorandum, 6 January, 1983; and AAMC guidelines in *Journal of Medical Education*, *op. cit.*, pp. 895–902.

[120] Wigodsky, Herman S., "Fraud and Misrepresentation in Research—Whose Responsibility?", *IRB* [Institutional Review Board]: *A Review of Human Subjects Research*, VI (March–April 1984), p. 4.

[121] *Ibid.*, p. 5.

[122] Brandt, E. T., *op. cit.*

observed by science-journalists. A meeting of minds—scientific and journalistic—could fashion a set of commonly respected principles of reporting. Such discussions of the criteria for intelligible translation from the scientific to the lay press, for finding informed experts—and not just those who are eminent—and for establishing harmonious relationships between the two professions, might increase the understanding of what "fairness" and "morality" mean to the journalist and the scientist.[123]

Conclusions and Prospects

Most scientific practice falls between squalor and greatness. At each stage in the process—research itself, reporting, professional publication and explanation to the wider public—the advancement of knowledge, the gratification of personal ambitions and the interests of patrons and users enter.

The profession of science weathers the assaults made on it by reaffirming its collective conscience regarding the ethical norms recommended and enforced to ensure compliance with the common goal of scientific discovery. It is necessary to re-examine the rules and procedures and to develop better ways of formulating and enforcing them.

Even a few cases of misconduct can afford significant opportunities for strengthening the moral practices of the sciences in which they occur. Scientists and their professional societies should undertake to refine their statements of these subtle expectations which resist codification. Such an improved statement of the ethics of scientific research would perhaps strengthen the traditions of proper scientific conduct.

[123] See Greenberg, Daniel, "Fraudian Analysis", *New Scientist*, VC (2 September, 1982), p. 643; David, Peter, "The System Defends Itself", *Nature*, CCCIII (2 June, 1983), p. 369; Steneck, Nicholas H., "Commentary: The University and Research Ethics", *Science, Technology & Human Values*, IX (Fall 1984), pp. 6–15; Edsall, John T., "Two Aspects of Scientific Responsibility", *Science*, CCXII (3 April, 1981), pp. 11–14.

Fraud in Science

Betrayers of the Truth. WILLIAM BROAD and NICHOLAS WADE. Simon and Schuster, New York, 1983. 256 pp. $14.95.

We are hearing a great deal these days about fraud in science. Since 1974, when Summerlin's partly black mice were found to owe their color to a felt-tipped pen, there have been several well-publicized instances of fraudulent research. Does this mean that there is more fraud, or is it simply that more attention is being focused on the subject?

I think both explanations are partly correct. There have always been instances of fraud in scientific research, prompted by such motivations as the desire for recognition, rivalry with another researcher, or a fanatical attachment to a theory. Now, however, there are new temptations. Before World War II, scientific research was engaged in by a relatively small number of devotees who generally had other means of earning a living, often teaching or, in the case of physicians, clinical practice. Now that the government has undertaken the major support of scientific research and the enterprise has grown enormously there are large numbers of individuals for whom research is an essential means of livelihood. Their advancement within their institutions and their claim on further government support are functions of their productivity. Since it is difficult and time-consuming to measure and agree on the quality of a scientist's output, a career is now evaluated in large part by the number of publications, with some adjustments made for the prestige of the journals in which they appear. No wonder that some career researchers are greatly tempted to cut corners in order to produce an acceptable manuscript. This is the new incentive to commit fraud; even without it, of course, the likelihood that fraud will occur has increased by virtue of the increased numbers of researchers.

At the same time that large-scale government funding of scientific research has attracted careerists to science, it has also led to greater public attention to fraud when it occurs. After all, since the public is paying for science, it has an interest in knowing how its money is spent. The media respond to the proprietary interest of the public and also to its appetites. Fraud is a sensational story, and scientific fraud has the added dramatic element of a fall from grace of privileged members of society whose stock in trade is truth. For the media the irony is compelling.

So it is that William Broad and Nicholas Wade have entitled their book on fraud in science *Betrayers of the Truth*. This is a short, excited book that chronicles selected episodes of cheating in science—from Ptolemy to Darsee—and offers them as support for three theses: (i) that fraud is common in science, and always has been; (ii) that the faith of scientists in their own methods makes them especially vulnerable to fraud; and (iii) that the response of the scientific community to instances of fraud has been inexcusably feeble. The nucleus of the book consists of the accounts of several recent instances of fraud in science, all occurring within the last decade. Five of these (the cases of Alsabti, Long, Soman, Spector, and Darsee) have already been reported in considerable detail by Broad or by Wade, mainly in *Science*. All five researchers were trying to succeed in a system that demands a constant and high rate of publication and they cheated to do it. The invitation to fraud inherent in the relentless pressure to publish is clear to Broad and Wade, and they also see the dangers in the "research mills" that in part grow out of this pressure—the large laboratories, headed by a senior researcher who attracts the grant money but grows increasingly remote from the actual work.

The stories of the recent cases of fraud occupy about half of *Betrayers of the Truth*. The remainder of the book consists of extravagant ruminations on the nature of science and an effort to give the subject historical sweep by recounting

episodes of fraud in science from earliest times. All of this is offered in support of the contentions that fraud in science is frequent and that it somehow grows out of the very nature of science.

When Broad and Wade leave off reporting and attempt this sort of analysis, they lose their way. Much of this material is superficial and pointless, and some of it is, I think, simply wrong. If Ptolemy, Galileo, and Newton were cheats, what are we to make of that? It certainly tells us nothing about the frequency of fraud in science. The point of these sketchy anecdotes, then, is not always clear, and where it is, it is often forced, in large part because there is no allowance for the fact that the accepted purposes and methods of science change with time. The Lysenko case is put to a particularly eccentric use, not to document the pitfalls of pursuing science in a risky and unpredictable political climate but to demonstrate "the limits to the ability of the scientific method to resist encroachment by nonscientific ideologies." This predisposition of science to intellectual corruption is a recurrent theme in the book. According to Broad and Wade, scientists' belief in their own objectivity and immunity from "rhetoric and propaganda" has the paradoxical effect of making them "the more susceptible" to these forms of argument and "strangely vulnerable to the unexpected." It is not clear whether Broad and Wade mean that scientists more easily fall into fraudulent practices or are unusually slow to recognize fraud in others, or both, but it sounds ominous, if murky. I will have more to say about this later.

How frequent is fraud in science? After considering the scanty evidence on the subject Broad and Wade conclude, "We would expect that for every case of major fraud that comes to light, a hundred or so go undetected. For each major fraud, perhaps a thousand minor fakeries are perpetrated." No basis for these numbers is given. This sort of calculation by expectation is surprising in such stern critics of scientists' estimates on the subject. It remains true, however, that the view held by some in the scientific community that fraud is in general not serious because it is rare and self-correcting is too sanguine. The truth is that no one knows whether fraud is rare or not or whether it will inevitably be found out. Though fraud with major scientific implications will probably be exposed by the scientific process in the long run (as in the Piltdown man hoax), this is not necessarily true of all fraud and certainly does not apply to the short run. Moreover, even rare instances of fraud damage the trust essential to the integrity of science.

What of the second thesis of Broad and Wade, that the conceit of scientists somehow renders them especially vulnerable to fraud? The weakest portions of the book are those designed to support this argument. Much here hangs on the authors' contention that scientists adhere to a myth of logic: "Researchers are imbued in their lengthy training with the notion that science is a realm of thought where logic and objectivity reign supreme." I don't think so. This is a straw man, meant to imply that scientists are unrealistic, hence easily deceived. In real life, scientists are nowhere near so breathless about their work. Instead, they are by and large very much aware of the temptation to go beyond the evidence and most of them resist it. In this regard it is curious that Broad and Wade imply that there was something reprehensible in the reluctance of scientists to embrace the theory of continental drift, despite its "intuitive plausibility," until there was "incontrovertible evidence of the sort that compelled even the blind to see." It is disconcerting to find a critical attitude toward the attachment of scientists to evidence in a book devoted to suggesting that they are too easily swayed by rhetoric and propaganda.

Perhaps the most serious error in the depiction of science in *Betrayers of the Truth* is the implication that scientists do not take into account the possibility of self-deception. This is not so. Much of the thrust of modern scientific methods is directed at this problem. As examples, the randomized controlled trial, the heavy reliance on statistical analysis, and indeed the very format of scientific reports are designed in large part to deal with the possibility of self-deception. Yet Broad and Wade represent the conventions of the scientific report—"as stylized as a sonnet"—as a sort of public relations trick "so as to give the appearance of objectivity." They favor a much freer form, one that would permit the scientist to describe "the excitement of discovery, the false leads, the hopes and disappointments, or even the path of thinking that may have led him through the various steps of his experiment." This idea conveys little understanding of the purpose of the scientific report. One reason for the impersonal format is, as Broad and Wade should know, to require the researcher to confront his or her data. In short, it is to reduce the potential for obfuscation, bias, and, I would add, tedium that would be introduced by an account of the author's agonies and ecstasies.

Broad and Wade are on their surest ground in arguing their third thesis, that the responses of the scientific community to the recent cases of fraud were inadequate. The responses varied somewhat, but in general were marked by confusion and a horror of "going public." They were also slow. In the Soman case, no attempt was made to evaluate the extent of Soman's fraud until a year after he was known to be guilty of plagiarism and accused of having fabricated data. Even where institutions and individuals dealt with fraud swiftly, they did so quietly, with the aim of removing the guilty researcher from the particular institution. There was little sense of urgency in notifying journals to which the guilty researcher had submitted manuscripts, other collaborators who might not know of the fraud, and other institutions where the investigator had worked earlier. It was almost as though the concern was more with avoiding the taint of fraud than with preventing its propagation in the scientific literature.

Broad and Wade are correct, then, in seeing the responses as inadequate. Their tone of outrage, however, is inappropriate and does not allow for the complexities and uncertainties of dealing with a situation as it unfolds, rather than in hindsight. Much of the disorganization and delay simply reflected the fact that the scientific community had never developed mechanisms for dealing with fraud. Without established procedures, individuals who are faced with the possibility of fraud in their midst are subject to a number of concerns that make it difficult to act. These include uncertainties about due process and the legal ramifications of making an accusation of fraud, as well as the desire to protect the reputation and funding of the laboratory and the institution. Furthermore, the recent cases were handled not only on an ad hoc basis but usually by people close to the guilty researcher and to the work, who were therefore subject to bias.

Because of these problems, several academic medical centers, including Harvard and Yale, have recently moved to establish uniform guidelines for dealing with accusations of fraud and assessing guilt. The Association of American Medical Colleges appointed an Ad Hoc Committee on the Maintenance of High Ethical Standards in the Conduct of Research, which issued its guidelines last year. Such efforts are important and necessary steps. We also need to examine systematically ways to modify the incentives to publish, so that the emphasis on quantity is reduced.

It would be a pity if the scientific community allowed the misconceptions in a book such as *Betrayers of the Truth* to deflect it from the important point that we must take fraud seriously and continue our efforts to develop mechanisms for dealing with it effectively.

MARCIA ANGELL
New England Journal of Medicine,
Boston, Massachusetts 02115

Trouble at t'lab

Walter Gratzer

Betrayers of the Truth: Fraud and Deceit in the Halls of Science.
By William Broad and Nicholas Wade.
Simon & Schuster: 1983. Pp.256. $14.95.

STORIES of heroic deeds in the laboratory act on the reading public like chloroform. Science becomes news when a microbiology technician catches bubonic plague or a professor is set on by a rampaging mob of anti-vivisectionists. Scientific fraud is a subject that has featured widely in science reporting of late, helped by the ill-suppressed embarrassment and indignation of the scientific establishment at the intrusion of the outside word into a private grief. The description in the opening chapter of *Betrayers of the Truth* of the inept attempts of a group of scientific mandarins, brought before a congressional committee on the subject, to get the toothpaste back in the tube is sheer bliss.

William Broad and Nicholas Wade have put together a most captivating book. We are most of us *voyeurs*, and nothing engenders so much prurient fascination as a gross and indecent violation of the rules by which we are expected to regulate our working lives. Throw in the accompanying discomfiture of stout and dignified parties in positions of authority, and you have all the components of what makes news on a wider stage.

Broad and Wade's purpose, however, is more than just to entertain, for they believe that a close study of the pathology of science will reveal aspects of its normal function that might otherwise remain hidden. Their thesis is that the scientific process is no more informed by reason and objectivity than are other human occupations, that it owes more than is generally conceded to cultural prejudice and rhetorical convention, and that even the best scientists employ in the pursuit of truth a mixture of motives and methods, many of which may not bear close scrutiny. This is not of course the first time that propositions of this nature have been expounded, but Broad and Wade go on to argue that the conventional wisdom about how science works is shown up for a shallow and meretricious conflation by the response of the system to fraud. They are concerned particularly to refute the often heard assertion that falsehood is rapidly eliminated through its failure to meet the criterion for all publishable work, namely that it must be capable of reproduction in other laboratories. In most of the case histories given in the book, the fraud was indeed not exposed by this mechanism. Most commonly it was uncovered accidentally, but more often, it seems to me, than Broad and Wade allow, a gamey smell issuing from published descriptions of the results was picked up by sensitive noses early enough to save expenditure of useless effort.

Leaving aside a number of now celebrated (though still perplexing) historical cases, in which discoverers of great truths (Ptolemy, Newton, Mendel and more recently it seems R.A. Millikan) apparently felt impelled to give their experimental observations a furtive helping hand, the frauds in areas of major importance described by Broad and Wade — xenografting, for example, and the mechanism of malignant transformation by a viral gene — tended to be quickly detected. This has not been so in that vast undergrowth of science, in which the sole criterion of accomplishment is publication, and the products are phantom papers in the sense that they are probably read by no one at all and will never be cited. One of the most remarkable cases in the book concerns a young postdoc, who perpetrated a series of exceptionally cool frauds while working at Harvard Medical School, and during his two years there produced one hundred publications. Can anyone doubt that the outcome of such proceedings must at best be trivial, and that it scarcely matters whether conscious deception plays any part in them? One might argue that the same holds for much of the output in the "soft" sciences — prudence deters me from being too specific — in which all results tend to reduce to the basic form "some do, some don't".

Betrayers of the Truth concerns itself not only with fraud, but also with other forms of professional misconduct. Indeed the authors do not make much of a distinction between the scholarly sin against God and those against man such as plagiarism, nor do they exclude mere featherbrained credulity, such as emerged so startlingly during the brief reign of Uri Geller and the infant spoon-benders. There are many ways of defrauding one's fellow scientists, the most common being probably the several forms of skulduggery open to referees of papers and grant applications. There are practically no penalties for such abuses;

> Thou shalt not steal; an empty feat,
> When it's so lucrative to cheat.

Outright fraud is generally fatal if discovered, though even here there are exceptions. In the only case of which I have first-hand knowledge, the perpetrator became a professor and is still hard at it ten years after his first transgressions came to light. So much for the consequences of breaking the tribal taboos. Perhaps just one ritual sacrifice every few years is sufficient to expunge all our accumulated sins.

The question that Broad and Wade do not attempt to answer is how much outright fraud actually retards scientific progress. The answer, I would hazard, is not very much, even if, as they surmise, there is a great deal more of it about than is apparent — in accordance perhaps with the natural law that seven-eighths of nearly everything is invisible. Fraud no doubt reduces the signal-to-noise ratio a little, but its effect is surely trivial compared to that of honest error. However Broad and Wade argue that the incidence of fraud in modern science is a symptom of corruption at the core. It stems, they suggest, from the trend towards large and impersonal research groups, in the thrall of robber barons, whose power and rapacity is seldom matched by the intellectual authority needed to run a team of twenty or thirty junior research workers. It is undoubtedly all too common for the leader of such a group to exert relentless pressure on its members to produce or else forfeit his patronage, or, worse, to encourage them to compete with each other for his favours. Add to this the stress that is still laid in many quarters on numbers of publications, which to the ignorant remain a more tangible measure of virtue than quality, and you are left with a rich compost in which unhealthy seeds will surely grow. The actions of many of the malefactors in cases of fraud are usually not those of rational men, and in their revelations from the penitent's stool, they speak most often of intolerable stresses applied usually from above.

I would also hold, however, that in a system that ultimately depends on good faith, no one is proof against an ingeniously contrived fraud and a number of blameless and respected scientists have undoubtedly suffered. Nor indeed would one suggest that circumstances can ever altogether exculpate the miscreant. Why, when all is said, did they do it? I think it was Sutton, the bank-robber, who was asked why he was so given to holding up banks; "because" he replied "that's where the money is".

A final reflection: a disproportionate number of the grosser cases of deception described in the book occurred in clinical departments of medical schools, where the exhortations to get ahead by way of publications appear often to be more insistent, and the critical standards consequently lower. There is also, by Broad and Wade's account, a marked reluctance to root out fraud when it occurs and impose penalties to match the offence. And what has happened, you may ask, to the egregious young men, who led their elders by the nose at Harvard, Yale and a clutch of lesser medical schools? Why, they have been sentenced to forfeiture of federal grants and condemned to drink the water of poverty and eat the bread of affliction in medical practice for a while. Dr Johnson said that a man is seldom so harmlessly occupied as when he is making money. Were I a patient, I think I would sooner these enterprising citizens were back in the laboratory accumulating inflated academic capital than wondering how they might turn me into an interesting publication. As an excited professor of medicine once said to me in a gratifying moment of self-revelation: "this patient is definitely publishable". □

Walter Gratzer is in the MRC Biophysics Unit, King's College, University of London.

Policies for Responding to Allegations of Fraud in Research

PENELOPE J. GREENE, JANE S. DURCH, WENDY HORWITZ AND VALWYN S. HOOPER

As a result of several cases of fraud in biomedical research, which have attracted much public attention, scientists and scientific organisations have been discussing what can or should be done when such charges arise.

Two major projects have touched on this problem: The Project on Professional Ethics of the American Association for the Advancement of Science in 1979 and 1980[1] and a workshop on "whistle-blowing" in biomedical research in 1981.[2] The Subcommittee on Investigations and Oversight of the Committee on Science and Technology of the United States House of Representatives held hearings in 1981 to receive testimony from prominent scientists and moral philosophers on dangers to the integrity of science.[3] The hearings showed that research institutions generally lacked any policies for dealing with allegations of fraud in research; they also showed that scientists were reluctant to have scientific work subjected to public scrutiny. The late Philip Handler, then president of the National Academy of Sciences, said in testimony: "I will admit with you, sir, the absence of any sense of what due process should be when some suspicion is raised. We have never adopted standardized procedures of any kind to deal with these isolated events."[4] When asked about the proper locus of responsibility for dealing with misconduct in research, he replied: "We would much prefer that the institutions within which science is conducted be responsible for their own affairs."[5]

Nevertheless, wide discussion is now under way. Recently Professors Warren Schmaus[6] and Harriet Zuckerman[7] have debated the difficulty of distinguishing error in research from fraud in research, and the sanctions

[1] Chalk, R., Frankel, M. S. and Chafer, S. B., *AAAS Professional Ethics Project* (Washington, DC: AAAS 1981).

[2] Swazey, J. P. and Scher, S. R. (eds), *Whistleblowing in Biomedical Research* (Washington, DC: US Government Printing Office, 1981). This workshop was sponsored by the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research; the American Association for the Advancement of Science, Committee on Scientific Freedom and Responsibility; and Medicine in the Public Interest.

[3] *Hearings before the Subcommittee on Investigations and Oversight of the Committee on Science and Technology*, United States House of Representatives (Washington, DC: US Government Printing Office, 1981).

[4] *Ibid.*

[5] *Ibid.*

[6] Schmaus, W., "Fraud and the Norms of Science", *Science, Technology and Human Values*, VIII (Autumn 1983), pp. 12–22.

[7] Zuckerman, H., "Norms and Deviant Behavior in Science", *Science, Technology and Human Values*, IX (Winter 1984), pp. 7–13.

appropriate to the latter. A number of research institutions and professional associations have begun to formulate policies.

The Inquiry

From late 1982 to early 1984, we conducted a survey of institutional policies regarding allegations of fraud in research.

Because so little was known about the number of institutions with such policies or the substance of the policies, we distributed an open-ended questionnaire[8] which would permit those to whom it was addressed to respond in detail.

In November 1982 the questionnaire was sent to 747 academic institutions and hospitals in the United States.[9] It was sent to the 118 members of the Association of American Medical Colleges, the 423 members of the Council of Teaching Hospitals, the graduate divisions of 183 major universities which grant doctorates, and to the 23 members of the Association of Schools of Public Health. The questionnaires were directed to the presidents or deans of these institutions. The responses, often handwritten, were returned by these officials or by their senior associates such as the vice-provost for research or an assistant chief of staff. In February and May of 1983, we wrote again to officials of institutions which had not returned the completed questionnaire. A total of 493 completed questionnaires were returned, a response rate of 66 per cent. The respondents included 94 of the medical schools, representing 80 per cent of the total; 255, or 60 per cent, of the teaching hospitals; 128, or 70 per cent, of the graduate schools of universities; and 16, or 70 per cent, of the schools of public health.

Results of the Inquiry

Of the 493 institutions responding, 116 had written rules for dealing with allegations of fraud; 124 had none at all and no plans to formulate any; the remainder were engaged in formulating such rules. There were fundamental differences in attitude towards the need for such rules. Typical of the attitude of one group of institutions was the statement: "Issues of research misconduct are likely to be so infrequent in their appearance and so difficult in their form that they do not lend themselves to being built into a code or

[8] Thirteen pre-test questionnaires were sent to senior academic administrators throughout the United States. Responses led to minor revisions and produced unsolicited reports of two episodes of fraud in research. In both instances, the entire proceeding was handled quietly within the department, and few persons knew of the problem or the reason for resignation.

[9] These packets included a cover letter explaining the survey, a stamped, pre-addressed acknowledgement postcard, the four-page questionnaire, and a stamped, pre-addressed return envelope. The cover letter was signed by Dr David Hamburg, then director of the division of health policy research and education of Harvard University, Professor Frederick Mosteller, chairman of the health science policy working group, and of the department of health policy and management, Harvard School of Public Health, and Penelope J. Greene, then Josiah Macy fellow in the division of health policy research and education of Harvard University.

guidelines." Another group expressed a very different view, represented by the following statement:

The impetus for this document [their policy] is the ample evidence from other institutions demonstrating that the absence of policies and procedures for dealing with allegations of scientific fraud and dishonesty has resulted in both unfortunate delays in resolving the matter and embarrassment to the involved parties and institutions.

Of the 116 institutions which had explicit and written policies, 38 had rules to deal specifically with fraud; 44 had broad rules governing the conduct of members of their staff, explicitly mentioning fraud in research; and 13 had broad rules, incorporated into faculty handbooks, personnel manuals, and general procedures for dealing with misconduct, which did not explicitly mention fraud in research. Among the remaining 21 institutions with written policies, the procedure by which charges of fraud would be handled was unclear, since some institutions with written policies described their policies rather than supplying the document containing the details (Table I).

One group of institutions had adopted policies, either written or unwritten, which stipulated a sequence of formal procedures. For example, one institution described its approach as follows:

Unwritten guidelines exist to the extent that the few instances in which misconduct has been alleged have been handled as in the following manner. The allegation has been presented to a member of the faculty senior to the individual involved—the division head or department chairman. That person would assemble enough facts to present the situation to the dean, who in turn has appointed an impartial committee to review the charges and recommend action. They have been instructed that any investigation must be conducted in compliance with rules of due process established in the faculty employment contract. If the review committee were to recommend disciplinary action, the situation would be reported to the University President for his concurrence.

TABLE I

Procedures Mentioned by Respondents to Questionnaire[a]

	By 116 with written policies		By the other 377		By all 493 respondents	
	(no.)	(per cent)	(no.)	(per cent)	(no.)	(per cent)
Informal resolution	31	27	12	3	43	9
Discussion	34	29	14	4	48	10
Investigation	80	69	130	35	210	43
Arbitration or mediation	21	18	2	1	23	5
Hearings	51	44	28	7	79	16
Appeals	41	35	22	6	63	13
Sanctions—specified	65	56	36	10	101	21
—unspecified	18	16	31	8	49	10

[a] Institutions may, of course, mention more than one procedure or none at all. Therefore totals may come to more or less than 100 per cent.

At other institutions, however, policies have been adopted which fail to recognise explicitly basic requirements of due process and administrative accountability. For example, one respondent stated, "We have had experience with research fraud on our campus. It was handled entirely at the laboratory level—from detection to dismissal . . .". Sixteen respondents simply reported some variant of doing something "suitable", e.g., "We would take appropriate action"; "We would handle it in a sensible and prompt manner"; "[We] would take the necessary steps"; or "We would do the right thing".

Of the 116 institutions with written policies, only 51, or 44 per cent, provide for hearings, and only 41, or 35 per cent, specify appeals procedures.

Two types of committees have been established to apply these policies: standing committees, whether or not designated to deal specifically with fraud, which always have a known composition; and *ad hoc* committees which vary in the degree of flexibility permitted in the appointment of members. An example of an *ad hoc* committee which has "membership specified" (Table II) is one composed of all departmental chairmen and the dean. In comparison, an example of an *ad hoc* committee of "partially specified" membership is one consisting of a limited number of departmental chairmen selected at the discretion of the dean. Some institutions state only that "a committee would be appointed".

TABLE II
Structure of Investigating Committees Cited by Respondents to Questionnaire[a]

	For 116 with written policies		For the other 377		For all 493 respondents	
	(no.)	(per cent)	(no.)	(per cent)	(no.)	(per cent)
Ad hoc committee						
Membership specified	12	10	7	2	19	4
Membership partially specified	45	39	38	10	83	17
Membership unspecified	26	22	62	16	88	18
Standing committee						
Dealing specifically with fraud in research	11	10	4	1	15	3
General but not institutional review board[b]	59	51	122	32	181	37
Institutional review board	6	5	33	9	39	8

[a] Institutions may, of course, mention more then one committee or none at all. Therefore totals may come to more or less than 100 per cent.

[b] Institutional review boards are established to evaluate research involving human subjects.

The prior designation of committee membership may be important to assure fairness because it offers no latitude to select a committee which will be, or be thought to be, advantageous or disadvantageous to the person charged. As one respondent stated: "Formation of a standing committee minimizes conflict of interest, and prejudicial selection of *ad hoc* reviewers, and allows for accumulation of experience by this Committee in such matters." Only with written rules is it certain whether committee membership is indeed fully specified in advance and thus independent of the individuals about whom allegations of fraud have been made. Of the 116 institutions with written policies, 45, or 39 per cent, rely on *ad hoc* committees with "partially specified" membership at some stage of their procedures, and 26, or 22 per cent, utilise *ad hoc* committees of "unspecified" membership; 12, or 10 per cent, have *ad hoc* committees with membership fully specified.

Among institutions which utilise standing committees, there are divergent opinions as to the role of the institutional review board which ordinarily serves to approve—or disapprove—of proposals for research on human subjects. Thirty-nine respondents reported that the institutional review board might be called upon to deal with allegations of fraud in research. A few respondents stated the opposite, namely that "the institutional review board should not be used for the investigation of a charge of misconduct".

Eight Issues

We identified eight salient and controversial issues pertaining to problems raised by fraud in research and institutional responses to allegations of such conduct. These are: the definition of fraud in research; the need for policies; distribution of policies; participants in review procedures; notification of supporting institutions and journals; the appropriateness of publicity; legal and quasi-legal considerations; and the effectiveness of peer review in detecting and deterring fraud.

There are major disagreements about all the issues among those who make policies for their institutions. The diversity of opinion among respondents arises in part from different views of the relative advantages and disadvantages of alternative policies. For example, the advantages of having policies formulated in detail are weighed against the value of flexibility in treating the unique aspects of each case. Some differences of opinion are functions of the fierce determination of some scientists to assert the moral probity of all scientists, while others acknowledge that not all scientific work is so conducted.

Definition of fraud in research: The definition of fraud in research is ambiguous and controversial. One of the most difficult questions is where to set the boundary between acceptable and unacceptable conduct. Many respondents distinguished between cases of deliberate misrepresentation,

falsification or manipulation of data on the one hand, and "genuine error" on the other. Their major concern is where the line should be drawn between fraud and negligence. Most scientists agree that fabrication, falsification and misrepresentation of data, along with plagiarism, should be regarded as fraud—but it remains a problem as to how errors attributable to careless record-keeping or unconscious bias are to be treated. Some respondents reported that, in practice, a distinction between sloppy record-keeping and conscious manipulation of data is difficult to make. How can it be decided whether an error was inadvertent or deliberate?

Some respondents offered clear-cut statements defining fraud in research. For example: "Fraud in research may be defined as the intentional fabrication, falsification, or stealing of data. The reporting of data that has been fabricated, falsified, or stolen is a separate, but equally odious offense." Another respondent, however, expressed concern that important distinctions would prove difficult to apply: "I would predict that the distinction between unconscious bias, conscious [improvement] of data, and the obvious manipulation of data, or fraud, will be found to be a difficult process." A third seemed to see no difficulty in definition: "I wonder whether specific guidelines regarding research fraud are necessary; after all, fraud is fraud and deceit is deceit."

Desirability of a policy: Is there a compelling need to establish a set of rules prior to any allegation of fraud?

In arguing for such rules, some respondents referred to the problems encountered by institutions as a result of a lack of an established policy. However, other respondents thought that no general policy was needed because each case should be dealt with on its own merits. There was a great diversity of views on this problem. We offer some of them:

"We will decide on a plan of action when a problem exists."

"Neither group [responsible for reviewing the need for rules] views the problem locally with any particular urgency to develop formal guidelines for dealing with it. The local attitude can best be described as, 'if it isn't broken, don't fix it.' "

"There is no excuse or reason to deal gently with a justified allegation of research misconduct, but circumstances differ so widely from case to case that we doubt the feasibility of developing uniform procedures."

"Like so many other institutions, we have been tardy in setting up official mechanisms for handling this problem because of what might be called a form of psychic numbing. It never occurred to us that it could happen to us. . . ." [It did. P. Greene *et al.*]

"My own feeling is rather strongly that we should not trouble to develop any such specific guidelines; my assumption is that were a case of fraud to appear, we would handle it in a sensible and prompt manner. If this proves, in the event, to be false (so that we make a mess of things the way some other institutions have already done!), then I would think of mounting the operation and taking faculty time to put in place some structure."

Several respondents gave specific reasons for having promulgated a policy, including an experience of fraud in research, the publicity given to

fraud in research elsewhere, the support of basic principles of scientific conduct, and awareness of the issues aroused by our questionnaire. Some institutions adopted the policies of other institutions or organisations. The set of recommendations published in June 1982, by the Association of American Medical Colleges (AAMC)[10] was the most commonly invoked: 56 institutions reported drawing on the recommendations of that association. These included 24 medical schools, 20 teaching hospitals, eight universities, and four schools of public health.

Distribution of statement of policy: How are members of the research staff to know what the rules are? Should each research worker receive a written copy of the institution's statement of policy? If rules are unwritten, or are not distributed, how will a member of the staff know what procedure is to be used in the event that he or she wishes to make an allegation or is accused of misconduct?

Respondents expressed widely differing views on the distribution of such information to the research staff:

"It is important that institutional rules and guidelines about the conduct of such research are widely disseminated to the investigators so that affected parties are aware of the norms that apply to them, that violations of the rules will be investigated and that sanctions, where appropriate, will be imposed."

"Individual researchers have been advised that guidelines have been adopted and that they are available for review if they so desire."

"[Distribution is] not necessary. Our policy is well known."

Of the 116 institutions reporting formal written rules, 50 stated that these policies are distributed to individual staff-members and 14 stated that they are not.

Participants: Respondents disagreed sharply about whether individuals in direct contact with the allegedly fraudulent research or with the accused should be members of the investigative or hearing body. One respondent declared that persons who participated in any preliminary investigations of the allegation, or who are members of the accused person's department, should not be members of the hearing board. In contrast, other respondents stated that the chairman of the accused person's department would co-ordinate the investigation, or that the investigating committee would include the immediate supervisor of the accused person and the head of the department as well as members of the institution from outside the department. One respondent stated that the accused individual should participate, along with the institution's director and senior administrative officers, in drawing up plans for the inquiry.

Respondents also disagreed about the participation of "outsiders". One thought that the matter should be kept as closely within the institution as possible; another that before the conclusion of the preliminary

[10] *The Maintenance of High Ethical Standards in the Conduct of Research* (Association of American Medical Colleges, 1982).

investigation, the dean should appoint three research workers or administrators from other institutions to review the findings and recommendations of the preliminary committee; a third respondent proposed that all members of the investigative committee should be from outside the institution where the alleged fraud was committed.

Fifty of the responding institutions, including 32 with written statements of policy, said that their policies made some provision for outsiders to participate in the proceedings.

Notification: Should grant-awarding agencies and journals be informed of an allegation of fraudulent research which they have supported or published? If so, at what point in the process should that be done? Should institutions with which the accused individual has been previously affiliated be notified, and when?

Respondents took a variety of positions on this issue. One said that grant-awarding agencies and editors of journals would be notified of charges or proceedings when the investigating or hearing committee had completed a preliminary investigation and had concluded that "formal" proceedings were required. Another stated that grant-awarding bodies and journals would be notified only when it has been formally determined that fraud had been committed and all appeals have been completed. One respondent said that his institution would adhere to the grant-awarding body's policies regarding notification. For federally supported research, one respondent's policy is, "Where a federal agency has sponsored the research, [we would] notify the agency that an investigation is in progress, without identifying the individual or the grant involved unless otherwise required by law."

But another respondent reported an instance in which an allegation of fraud in research was brought to the attention of a federal grant-awarding body, only to be told that the agency "would not act unless someone else presented them with the results of investigation".

Sixty of the institutions which responded to the questionnaire, including 38 with written policies, stated that their procedures include notification to the body supporting the research or the journals which have published or plan to publish the results that an investigation into fraud is under way. One respondent said explicitly that the policy of his institution did not include such notification.

Publicity: Respondents disagreed about whether, how, when and to whom public announcement of an alleged fraud in research should be made. Some policies reflect a belief that wide publicity may act as a deterrent, while others emphasise the policy-maker's desire to protect the reputation of the institution.

Some respondents stated that hearings and their outcomes are not normally given publicity. At one institution, after a final decision in a case is made, the president decides whether or not to release information about it; he may disclose the information sooner if he deems it necessary. The

respondent for this institution expressed concern about the harm that might be done by releasing the information as well as from not releasing it. Another respondent said: "In consultation with legal counsel, institutional administrators may release appropriate information about the incident, particularly when public funds were used in supporting the research." Several respondents stated that disclosure is important for maintaining the integrity of research.

Procedural and Legal Issues

Requirements for acceptance of accusations: Some respondents reported that only written, signed charges are considered. At other institutions, anonymous accusations are accepted. One institution would proceed immediately with a full investigation in response to any written, signed charge; with an unwritten accusation, it would first consider the nature of the charge, the need for immediate action, the effect of the investigation on the accused, and the credibility of the person making the accusation.

Right to counsel: Many of the policies described by respondents permit legal counsel to act on behalf of the accused person. There are several different approaches to this point. One respondent said that in his institution:

> Attorneys may not appear with or on behalf of the grievant, witnesses, or the University in proceedings before the Senate Committee on Professional Ethics and Academic Freedom. It is assumed that either party may rely on legal counsel in the preparation of any documents or the collection of any evidence to be presented to the Committee.

Another said: "The faculty member has a right to be represented by any person as academic advisor and by counsel of his choice at all stages of the proceedings."

At one institution, both the accused and accuser may be represented by counsel, and an attorney may sit with the committee as hearing officer to "make all decisions on evidence and procedures, but not vote on the substantive merits of the request". Another respondent stated that in the event that his institution brought charges, it would "probably require a quasi-legal process—defense attorney, etc.".

Eighty-two of the respondents' institutions, including 45 with written policies, allow for legal counsel in some fashion.

Due process: The institutions differed sharply in the degree of attention they paid to assuring due process for accused individuals. Some respondents emphasised the importance of this issue. Two examples: "Any investigation must be conducted in compliance with rules of due process established in the faculty contract." "Because procedures are not more detailed, we would need to proceed carefully to ensure the rights of the accused and to provide for due process."

Other respondents failed to deal explicitly with adherence to rules of due

process. One group had policies so general as to be unclear whether due process is provided for. A comment from one respondent demonstrates an interest in fair and reasonable proceedings but leaves doubts about successfully ensuring due process: "The hearing is not conducted according to legal rules of evidence and witnesses. Any oral or documentary evidence may be received so long as it is of the sort on which responsible people are accustomed to rely in the conduct of serious affairs."

In a recent court case involving Princeton University's application of sanctions for plagiarism, the court made a distinction between obligations of due process of private institutions and those of public institutions such as state universities or municipal hospitals.[11]

Defamation of character and invasion of privacy: "Defamation of character" is the publication of untrue facts about a person, either in writing (libel) or orally (slander). An individual's "privacy" is invaded when true but private and embarrassing facts are published or if physical privacy is intruded upon, e.g., unauthorised entry into office or records. The possibility of defamation is an important consideration with respect to fraud in research. There is always the possibility that a charge of defamation of character or invasion of privacy will be lodged against the person making the accusation or against those engaged in the subsequent proceedings. Although it is not clear whether the statement below refers specifically to defamation or invasion of privacy, it does suggest that at least one institution has given thought to the question of liability of the person making the accusation: "The Vice-President for Health Affairs will assure sufficient protection against personal liability for actions taken during the review process for extramural as well as intramural members of the [review] board."

[11] 453 A.2d 279 (N.J.Super.Ch 1982); 453 A.2d 263 (N.J.Super.A.D. 1982). Due process has both a general meaning, pertaining to whether an individual is accorded the rights prescribed for the class of cases to which the individual belongs (Black, H. C., *Black's Law Dictionary*, 5th edn, St Paul, Minn.: West Publishing Co., 1979, p. 449) and a strict constitutional meaning, contained in the Sixth Amendment to the United States Constitution, pertaining to the rights of the accused in criminal prosecutions, including speedy trial, an impartial jury, confronting witnesses, having "the assistance of Counsel for his defense", etc. The relevance of the strict constitutional definition requires an analysis of whether an institution is one required under the Constitution to provide due process because of its governmental or quasi-governmental status. An example of the difference in the applications of these two meanings of due process is contained in the case of *Napolitano* v. *The Trustees of Princeton University, et al.*, in which a student penalised for plagiarism brought a legal action against the university. In ruling on an earlier decision of the superior court to dismiss the case, the appellate court held that the issue properly depended on "whether Princeton substantially complied with its own regulations . . . whether Princeton's decision was supported by the evidence . . . and whether the penalty imposed was within Princeton's authority to impose". (453 A.2d 263 [N.J.Super.A.D. 1982]). The appellate court judges distinguished the Napolitano case from the case of a medical student dismissed from a state medical school (*Greenhill* v. *Bailey*, 519 F.2ds). The Napolitano court observed that "judicial intervention is required when a state educational institution acts to deprive an individual of significant interest in either liberty or property".

Deterrence and the role of peer review: Many of those testifying at the congressional hearings in 1981 focused on the self-correcting nature of science as manifested in the system of peer review—through both the assessment of papers submitted to journals and the assessment of applications for financial support. For example, during these hearings, Dr Handler stated:

> The matter of falsification of data, I contend, need not be a matter of general societal concern. It is, rather, a relatively small matter which is generated in and is normally effectively managed by that smaller segment of the larger society which is the scientific community. This occurs in a system that operates in a highly effective democratic, self-correcting mode—the "peer review" system . . .

Dr Patricia Woolf, at the same congressional hearings and in a subsequent paper,[12] emphasised the role of regular peer review as a mechanism for deterring and discovering fraud in research. She referred also to the then recent establishment by the department of medicine at Yale University of institutional procedures to deal with allegations of fraud in research. The arguments advanced in 1981 by Drs Handler and Woolf show how very recent in origin is interest growing today in special institutional policies pertaining to fraud. The recent change in attitude on the part of some within science was expressed in a statement made in 1983 by Dr Arnold Relman, editor of *The New England Journal of Medicine*:

> There is a widespread opinion that science is self-correcting and that fraud will be exposed relatively promptly, as other scientists attempt to repeat the original work. That view was suggested in the recent report of an Association of American Medical Colleges Ad Hoc Committee on the Maintenance of High Ethical Standards in the Conduct of Research: "The principal deterrent in research fraud," declared the committee, "is the overwhelming probability that fraudulent data will be detected soon after their presentation." As a member of that committee, I now feel that this statement is too sanguine.[13]

Nevertheless, responses to our survey indicate a continued widespread belief in the reliability of traditional processes of self-correction in science. As a result, widespread scepticism persists concerning the need for institutional policies pertaining to fraud:

> "It is recognized that there is a near certainty that falsified, fabricated, or plagiarized data will be detected."
>
> "We would note that in our small medical school we have an almost 'village intimacy' (good and bad). Colleagues tend to know who is rough on glassware—far less [sic] who is bending data."
>
> "We expect that our continuing collegial monitoring of research would minimize the likelihood of a need for a full-blown 'investigation'."
>
> "The academic excellence and climate promoted by the peer researchers for

[12] Woolf, P., "Fraud in Science: How Much, How Serious?", *Hastings Center Report*, XI (October 1981), pp. 9–14.

[13] Relman, A. S., "Lessons from the Darsee Affair", *New England Journal of Medicine*, CCCVIII (June 1983), pp. 1,415–1,417.

quality research efforts and meaningful research act as their own deterrents to misconduct on the part of researchers."

One respondent stated that all research is reviewed annually by a research committee. Another institution relies on the peer-review process of journals and the regular presentation of research to peers.

Other respondents, however, expressed a belief in the deterrent value of explicit institutional policies regarding research fraud:

"The establishment of a well advertised procedure to deal with charges of fraud is considered a powerful deterrent . . ."

"It would be worthwhile to appoint a campus-wide committee to deal with the investigation of fraud should it occur. Although such a committee might never be used, having it in place would alert the campus community that you view such matters seriously."

A respondent, who had recently dealt with a documented case of fraud in federally supported research, made a suggestion which he admitted might be an extreme solution:

One way to diminish the frequency of cheating in research would be to have the position of the federal government clarified with regard to falsification of material placed in progress reports to the NIH, or published with credit to support by an NIH grant. If that were clearly a federal crime, what could be a relatively common event would diminish markedly in incidence.

Conclusion

Persons responsible for the formulation of policy within institutions need to consider a broad range of questions in developing policies to deal with allegations of fraud in research. These questions include the following: What constitutes scientific fraud? Should institutional policies include provisions for procedural steps to be followed in the event of an allegation of fraud? Must such provisions be written? If so, how should they be distributed? If not, how should they be publicised? Need the rules deal explicitly with fraud in research? Should committees which review allegations of fraud be standing committees to avoid any actual or apparent favouritism, or should they be formed *ad hoc* to allow flexibility for the consideration of each case on its own? What are appropriate sanctions, and should they be specified in advance? Should participants from outside the involved institution be included? Should grant-awarding agencies or journals, or both, be notified, and if so, at what point? Should information be released to the public, and if so, under what conditions? Should new mechanisms, supplementing regular peer review, be developed to deter and detect fraud in research? And finally, does an institution's policy deal adequately with the requirements of due process?

Our survey shows that vast differences of opinion exist as to whether institutional policies for responding to allegations of fraud in research are necessary, and how such policies should be implemented. At a conference in

November 1984, on "Institutional Responsibility for the Conduct of Research"[14] participants sharply disputed issues similar to those which were raised by respondents to our survey. The debate demonstrated the absence of consensus on these matters. In order to avoid the risk of precipitous decisions made under pressure of time, members of research institutions should try to resolve these complex questions before untoward events and legal necessity force them to do so.

[14] Sponsored by Public Responsibility in Medicine and Research, Boston, Massachusetts, 1–2 November, 1984.

II

POLICY FOR DEALING WITH FACULTY FRAUD IN RESEARCH

UNIVERSITY OF FLORIDA

I. Faculty Conduct in Research

A. *Institutional policy*

It is the policy of this institution that each individual faculty member is expected to maintain high ethical standards in the conduct and reporting of his/her academic research. Should allegations of fraud or misconduct in academic research occur, it is the duty of the faculty to respond in a fitting manner to resolve issues arising from such misconduct or fraud.

B. *Faculty responsibilities*

The faculty members of the University of Florida pledge to maintain ethical standards in the conduct and reporting of scientific research. The faculty recognizes that it has responsibilities for ethical conduct in research not only to the university but also to the community at large as well as to private and public institutions to which members of the faculty are bound by affiliation or contract.

II. Basic Principles Governing Fraud in Research

If allegations of fraud in research occur, a flexible response is required because the type, seriousness and technical nature of the alleged fraud may cover a broad spectrum. The faculty and its representatives will be guided by the following principles:

1. The rights of individual faculty members, students, and all employees of the University must be protected whether they be accused or whether they be the accusers.
2. The rules of "due process" shall govern any proceedings initiated in response to an allegation of fraud in research.
3. Confidentiality shall be maintained throughout an investigation to the extent that is possible in conducting a thorough investigation and that is consistent with the laws of the state of Florida.

III. Definition of Fraud

Fraud is defined as the intentional misrepresentation, concealment, "stealing" or non-disclosure of research data which is intended to deceive or in fact deceives another. An example of fraud would be intentional falsification of data "through selective reporting, including the omission of conflicting data."[1]

IV. Procedures for Dealing with Alleged Research Fraud

A. *Processing Initial Report of Fraud*

1. Allegations of misconduct or fraud in research should be brought forward by anyone involved in or aware of such misconduct.
2. Such allegations shall be put in writing, signed and submitted to the administrative officer to whom the accused reports, e.g., department chairman.
3. The department chairman (or equivalent administrative officer) shall immediately inform the dean of the college, in writing, of the allegation(s). The dean shall inform the vice president to whom he or she reports and the vice president for research of the existence of the allegations.
4. The dean shall also inform the accused faculty member and any collaborators in the research of the allegation(s) and the steps to be taken during the investigation. The accused individual will be afforded ample opportunity to respond to the allegation(s) and to provide additional information.
5. In the case of joint appointments and collaborative research projects, the dean of the college where the accused faculty member holds his/her primary appointment shall decide if other administrative units should be involved in the resolution of the allegation(s).
6. In the initial investigation, the dean may either (a) request the department chairman (or equivalent administrative officer) to conduct a preliminary inquiry of the allegation(s); (b) conduct such inquiry himself or

[1] *Report of the Association of American Universities*: Committee on the Integrity of Research.

herself to decide if it is appropriate to resolve the allegation(s) in an informal manner without invoking more formal administrative procedures; or (c) refer the allegation(s) directly to the committee on ethical standards (see Section IV, B1) for a thorough investigation.

7. The dean shall insure that an investigation, if any, is commenced within ten (10) days of notification of the allegation(s) unless appropriate justification exists for a delay. The individual accused of misconduct or the person making the accusations shall have the right to request that the allegation(s) be thoroughly investigated by the committee on ethical standards. The dean shall consider such request in his/her decision of whether to forward the allegation(s) to the committee on ethical standards.

8. If the dean chooses to have the department chairman or other appropriate administrator conduct a preliminary inquiry of the allegation(s), the department chairman shall conduct such inquiry and shall submit a written opinion to the dean (with copies to the vice president to whom the dean reports and the vice president for research). The written opinion shall indicate that the allegation(s) lack substance and no further action is recommended or that the allegation(s) have substance and further action is recommended in the matter. The department chairman shall submit his or her recommendation to the dean within a reasonable time of notification by the dean, not to exceed ninety (90) days without appropriate justification. After receiving the written report, the dean shall review the matter and submit a written recommendation within twenty–one (21) days to the vice president to whom he/she reports and the vice president for research that: (a) The allegation(s) lack substance and no further action is necessary; (b) the allegation(s) have substance and should be further investigated by the college's committee on ethical standards.

B. *Investigation of Reported Fraud that Appears Substantial*

1. The dean of the college in which the accused individual holds his/her primary appointment shall forward all substantial allegation(s) to the college's committee on ethical standards. This committee will be appointed by the dean of the college and shall consist of five faculty members from the college who have sufficient expertise about the subject matter under review. In addition, the dean shall appoint an attorney who will serve as a non-voting member of the committee to provide its members with legal advice and counsel.

2. The committee on ethical standards shall conduct a thorough investigation of the allegation(s). After completing its investigation, the committee shall submit its findings and recommendation(s), in writing, to the dean. The committee shall submit such findings and recommendations to the dean within a reasonable time of notification of such allegations, but not to exceed ninety (90) days with appropriate justification.

3. After receiving the written report from the committee, the dean shall review the matter and submit a written recommendation of action, within twenty-one (21) days, to the vice president to whom he or she reports and the vice president for research.

4. Throughout the investigation, the office of the dean and the committee on ethical standards shall maintain written records of all proceedings in

which they are involved during the resolution of the allegation(s). The accused faculty member shall be informed of any recommendation submitted to the dean and vice presidents.

5. In addition, the dean may, at his/her discretion, request the accused faculty member to withdraw all pending papers that have emanated from the alleged fraudulent research. The editors of the journals in which papers on this research are pending publication or have already appeared should be notified of the matter. In recognition of the responsibility of the faculty to the public and to the professional community at large, the dean may, after consultation with legal counsel, notify other appropriate individuals and institutions. Any public or private granting agencies sponsoring the research involved should be notified of the allegation(s). The accused faculty member shall, upon the dean's request, provide the dean with information which would enable the dean to make a determination as to the appropriate individuals or institutions to notify. It shall be within the dean's discretion to determine, in light of the facts of each case, when the appropriate individuals or institutions will be notified.

C. *Subsequent Action Following Completed Investigation*

1. Upon receipt of the dean's recommendation, the appropriate vice president shall review the matter and take such action as is deemed necessary. The accused faculty member shall be promptly informed of such action. Where formal charges are recommended by the dean, the vice president may confer with those involved and, if he or she concurs with the judgment of the dean that formal charges are required, shall formulate a formal recommendation for the president's consideration with advice from the vice president for academic affairs.

2. Appropriate disciplinary action should be considered in accordance with Florida Administrative Code Rules 6C1-7.10(1)(e), 6C1-7.36, 6C1-7.45, 6C1-7.46, and 6C1-7.48.

3. When appropriate, a recommendation shall be submitted by the dean or the vice president for academic affairs to the president. The president shall review the matter and may conduct further investigation if deemed necessary. The president shall take final action in the matter and shall notify the faculty member involved of the decision.

APPENDIX 7

GENERAL DISCUSSION OF RESEARCH PUBLICATION PRACTICES, ETHICS AND FRAUD

June 16, 1950, Vol. 111 SCIENCE 643

An Ethical Code for Scientists

Ward Pigman and Emmett B. Carmichael
University of Alabama, Medical-Dental Schools, Birmingham

A NEW PHENOMENON of our present-day society is the obviously important role played by science. Only a short time ago science was considered by many "practical" men as a plaything of inconsequential importance in contributing to the welfare of society. Although the significance of science was becoming more generally evident before World War II, this war demonstrated to the public in general and to legislators and businessmen in particular that science, especially basic science, is much more than a scholarly pursuit—that it is a vital force for the advancement or destruction of society. Science is now "big business." As a result, the scientist cannot and must not remain a scholarly recluse divorced from the remainder of society. His behavior and that of society toward him will greatly influence the progress of science and, to an increasing extent, that of society itself.

During its long period of development, science has evolved a code of professional tradition and ethics, largely in an unwritten form. This code, really the foundation of the scientific method in many of its aspects, has to a considerable extent been responsible for the achievements of science. Polanyi's (*12*) description of the effect of disregard for scientific traditions is applicable to many of our modern industrial and research organizations:

> Those who have visited the parts of the world where scientific life is just beginning, know of the backbreaking struggle that the lack of scientific tradition imposes on the pioneers. Here research work stagnates for lack of stimulus, there it runs wild in the absence of any proper directive influence. Unsound reputations grow like mushrooms: based on nothing but commonplace achievements, or even on mere empty boasts. Politics and business play havoc with appointments and the granting of subsidies for research. However rich the fund of local genius may be, such environment will fail to bring it to fruition.

The important achievements of science and its contributions to our civilization seem adequate proof of the basic validity of these traditions. On the other hand, conditions of scientific work have changed greatly, and obviously the traditions must be interpreted in terms of prevailing conditions. Science has emerged from a period in which the predominant effort was made by individuals, sometimes of almost an amateur status, to a period marked by the development of large research groups, many in the pursuit of research for profit. As a result, it is timely for the scientist to consider his professional traditions and to relate them in terms of the structure of modern scientific work.

These traditions are essentially an unwritten code of professional ethics. As pointed out by Leake (*9*), the term "professional ethics" as used generally includes the attitude of the individual scientist to society and to other scientists. It will be so used here. This concept of professional ethics inextricably involves social obligations, questions of etiquette, and adherence to accepted traditions. Claude Bernard (*1*) has contributed one of the better discussions of the ethical qualities needed in scientists, and the relationship of these qualities to the scientific method, although his remarks apply in the main to medicine and physiology.

Some of our professional organizations have established formal written codes of professional ethics.[1] In the medical field numerous papers and books have been written on the subject. One of the first extensive codifications was that of Percival (*9*) (1803), but the general precepts of Hippocrates (circa 500 B. C.) have modern acceptance. A major consideration at the first meeting of the American Medical Association in 1847 was the formulation of a code of professional ethics (*9*). The present code provides means for enforcement by its members. There has been some discussion of the professional responsibilities of industrial chemists (*6*) and a code has been proposed for this group (*7*). Scientific groups generally, however, have not formalized their traditions but have passed them on by example and by word of mouth as an informal part of the graduate student's training.

This failure of scientists as a group to consider ethics is revealed in the fact that *Chemical Abstracts*, since it was founded in 1907, listed only four refer-

[1] American Medical Association. *Principles of Medical Ethics* (*14*); American Dental Association. *Principles of Ethics* (*15*); American Association of University Professors. *Statements of Principle in Academic Freedom and Tenure* (*16*); American Institute of Chemical Engineers. *Constitution*, Article VIII (*17*); Engineers' Council for Professional Development. *Canons of Ethics for Engineers* (*18*). "The Geneticists' Manifesto" adopted at the Seventh International Congress of Genetics, held at Edinburgh in 1939. Published in *The Journal of Heredity*, September, 1939. Reprints distributed by American Genetics Association (*19*).

ences under ethics in its indices. It is true that professional codes at best can only express an ideal; their acceptance and application will depend upon the individual scientists. We believe, however, that the scientist's position in the world today makes it extremely important that his time-proved traditions be reconsidered in terms of modern circumstances and possibly written into a formal code. We believe that such an action would maintain the advance of science, increase its public support, and improve the professional relations of scientists. Improved professional relations would better morale and increase productivity among research men. Mills (*10*) has pointed out the social implications of ethical behavior in the distribution of research grants.

The planning of an ethical code for scientists should take into account first the scientist's general obligations as a member of society, and beyond that his special obligation as a scientist to protect society—here, there are many problems related to warfare, to the health and general well-being of mankind, and to nationalism versus internationalism. Such a code should preserve the scientist's ethical traditions and incorporate the scientific method. It should state the scientist's obligation to explain the nature and purposes of science, and the policies in dealing directly with the public. It should clarify the scientist's attitudes toward patents and secrecy restrictions. It should affirm the scientist's obligations to individuals—to his employer, his associates, other scientists, and his assistants and graduates—and scientists' obligation as a group to other professions.

We have merely indicated the scope of the problem. To deal with it fully in all its phases would require the efforts of scientists in many different fields of study and kinds of employment. Some of these phases have already received considerable attention. Because the results of atomic research have such unmistakable implications for society, attention has been paid to the scientist's attitude on the use of his discoveries, particularly for military purposes, and to the necessity of his being socially conscious (*3, 4, 5, 8, 10, 13*). Other phases of the problem have received little or no public consideration.

Many of the scientist's obligations are reciprocal in the sense that the scientist has grown to expect certain conditions for his work, and to a considerable extent these conditions affect the quality of his work. Sometimes his obligations are conflicting. He may at times be faced with the dilemma of obligations to his employer that conflict with obligations to the public as a whole. What should his attitude be when his employer's immediate interest causes harm to the general public? Suppose that his employer is a company that is pouring waste products into a stream and he knows that at a reasonable cost this pollution could be greatly minimized. Should he assume, as a lawyer does, that his primary obligation is to his client, and become an automatic defender of the company's position? Or should he consider that he has duties to society greater than those to the company?

A group of very pressing problems is presented in the application of traditions related to the authorship and publication of scientific researches. In the early days of science most articles carried the name of only one worker, whereas multiple authorship is now most common and sometimes ten or more persons may be involved. As a detailed example of the need for a code of professional ethics, we will discuss some of the problems involved in authorship.

General Obligations of Authors

Quality of Papers. Everyone will agree that scientific articles should be of good quality, should be original in content, and should describe all work in a reproducible fashion. These are fundamental requirements of the scientific method and yet most scientists would admit that many research articles are published that are deficient in some or all of these respects.

Claude Bernard (*2*) has described the importance of adequate details:

> In scientific investigation, minutiae of method are of the highest importance. The happy choice of an animal, an instrument constructed in some special way, one reagent used instead of another, may often suffice to solve the most abstract and lofty questions . . . . the greatest scientific truths are rooted in details of experimental investigation which form, as it were, the soil in which these truths develop.

Even casual inspection will show that many articles are not written so that the work can be repeated. Traditional procedure is often ignored in reporting new compounds; occasional articles will not give analyses of new compounds or the compounds will be poorly described so that their identity is questionable. Scientific journals lack space to print all the good material they receive today, and understandably urge authors to shorten their articles, but great care is needed to avoid eliminating important details.

Direct responsibility to prior work. The traditions of science demand that any report of scientific work must consider prior work, integrate it in the general subject, and cite proper references to it. Frequent violations of this principle must be familiar to all scientists. One of us has previously called attention to an instance of this type, particularly in relation to the naming of methods (*11*).

The basic concept behind this principle, even more fundamental than professional courtesy, is that fre-

quently the solution to a problem may already be in the literature and needless repetition is economic waste. Thorough literature searching can be defended from an economic standpoint alone. In order to speed the incorporation of new work into the general body of basic knowledge, each author has the responsibility of assisting in the integration of his work with that of previous workers.

To many scientists, establishment of priority for new discoveries is important, and organizations that seek patent protection for their work may set up involved and expensive procedures to establish the date of discovery. Is this a tradition that should be continued? Some scientific journals do not carry the date a manuscript was submitted, and few indicate whether essential changes of content have been made after that date. "Letters to the Editor" may require careful controls to prevent abuses.

Criticism and disagreement. The scientific method requires that all research work be open to critical examination and testing by researchers in the field. It also requires that dissenting theories and results be treated with tolerance, and not suppressed merely because they disagree with currently accepted ideas. Many scientists would add that mistakes and errors should be publicly acknowledged.

The widespread violation of these principles today is affecting not only the progress of science but our economy as well. Many commercial research organizations keep closed files of their researches as a matter of policy, in the belief that they will have an advantage over their competitors. Most of them do not realize that lack of criticism of the worker by qualified colleagues in his own field fosters the carrying out and perpetuation of poor or erroneous work, the continued employment and promotion of unqualified workers, and the perpetuation of poor research policies. Criticism by members of the worker's organization and by consultants is usually inadequate because of the influence of personal motives and lack of knowledge in the specific field. Objections to excessive secrecy in military research should take into account this principle as a primary consideration.

Classical examples of the value of scientific controversy are well known. When properly conducted, such debates lead to clarification and advancement of knowledge. But improperly conducted, they lead to enduring feuds, and because of this possibility, there is a tendency among editors of journals to suppress scientific polemics. A continuation or extension of this trend will be a severe blow to the scientific method. However, as stated by Wise (*13*)

> the research worker should not permit himself to become embittered or involved in useless polemics. . . . It simply means that his criticisms must be objective and that they must not descend to the plane of personalities. He must show that he is dealing with a set of data, not with an enemy.

Property rights of the scientist in his work. A currently controversial problem of the application of scientific tradition involves the rights of a researcher to his work. The decision to try to publish or reveal his research once was the sole right of the scientist. Now, with the investigator receiving financial support from others in most cases, the final decision is tending to fall on the provider of the funds. In the extreme case, what is there to prevent someone in authority from taking over the work of an associate and passing it off as his own work? What should an editor of a scientific journal do if he receives for publication a suitable manuscript from an established worker and simultaneously a letter from the supporting group saying that it should not be published? Should the supporting group be required to provide satisfactory and convincing grounds? By tradition and perhaps even by legal mandate, the rights of an artist to certain phases of the disposition of his work have been affirmed. Should these not apply equally to the scientist, whose application of his science is often an art?

At least one established graduate institution has the policy that all doctoral theses are published solely in the names of the individual graduate students. In certain instances, the idea was suggested by a member of the faculty, who carried out some preliminary work, supervised the principal research, drew or helped draw the conclusions, rewrote the thesis and wrote the final published version. Is this an example of acceptable ethics?

Publicity. Some scientists violently oppose general publicity and popularization of their work. Others seek publicity, and some even condone or support erroneous and misleading publicity. What should be the attitude of the scientist? Does he owe the public a duty to attempt to explain the purpose and significance of his work? Should chicanery and excessive or misleading publicity on the part of scientists and nonscientists be exposed as a function of scientific societies? It is of interest that the *Principles of Medical Ethics* includes a considerable discussion of the impropriety of advertising and publicity-seeking and that AMA members are required to advise the public against misrepresentation. A firm stand on this issue by scientists generally might be of considerable help in establishing the professional status of the scientist in the public mind.

Multiple Authorship

With the change in status of research, owing to its being produced not by independent individuals but by

several dependent workers or even large groups, the scientific tradition in respect to the etiquette of authorship needs reinterpretation or extension. The responsibilities involved in multiple authorship or group research must be analyzed.

"Senior" authorship and order of names. To many scientists, the order of the authors' names on a publication has a significance. Is this a tradition that should be preserved, clarified, and enforced, or is it an outmoded, unessential form of etiquette? In current publications, the application seems uncertain and haphazard. Should the concept of the "senior author" (the first one listed) be preserved? If so, should the senior author be the person highest in the administrative rank, the one who has done most of the laboratory work, the one who has written the paper, the one who furnished the original idea, or the one whose technical skill and thoughts have carried along the research?

Administrators and financial supporters. In publications, what consideration should be given to administrators and financial supporters? Some scientists might say that they should be indicated as authors only when their contribution to the actual solution of the problem has been substantial, continuous, and of a high level. Probably most scientists would agree that mere general administrative supervision of a project or even the suggestion of the original idea for the project is insufficient for an authorship. Certainly no one should be granted authorship of any type merely because he has seniority or is in charge of a laboratory. We cite as an example a man serving as technical liaison between a company and a research organization who insisted that his name be included as an author, before he would ask for supporting funds for the research, although his total contribution was limited to this action.

Graduate students and technical assistants. Criteria are necessary for assessing the role of graduate students and technical assistants in relation to authorship. Should not senior authorship for a graduate student be limited to those instances in which a real contribution, beyond adequate laboratory work, has been made? On the other hand, is it not the duty of the directing professor to encourage the student to his maximum performance, rather than use him as a laboratory assistant? If a technical assistant is to be given authorship of any type, more than an adequate performance of routine methods should be required of him.

Group projects. An example of the large group projects that characterize modern science is the penicillin research during World War II. Industrial organizations provide many more examples. Frequently there is no attempt on the part of administrators to set up the program so that the work of individual investigators is kept discrete. The improved quality of work resulting from the establishment of definite responsibility might be the basis for making a definite statement in regard to this problem. The interpretation of the scientific tradition in terms of modern group research is an extremely important and as yet unexplored field.

Preparation of manuscripts. The published paper is the final record of the finished research work, and the medium through which the information is made generally available and useful. With the present shortage of publication space, the preparation of the manuscript becomes more important than ever. Rigid adherence to established scientific traditions on the part of authors and editors becomes increasingly essential.

To many persons, the preparation of a research paper may seem to be a routine matter, but actually it requires a high order of skill and technical knowledge and an acquaintance with scientific traditions. In many researches the actual preparation of the manuscript, the integration of the findings with the prior related work, consideration of the significance of the data, and arrangements for publication may require a considerable portion of the time and skill required for the entire project. Possibly the actual preparation of the manuscript should be a factor in defining the responsibilities of the senior author. Laboratory workers without a good background of knowledge, and research administrators without close daily contact with the laboratory work and a thorough knowledge of the field probably should be discouraged from actual preparation of the manuscript. On the other hand, simple manuscript revision, in spite of the poor writing ability of many scientists, should not generally be made the basis for authorship of a research paper. Still another problem is determining the duties and responsibilities of the referees of scientific articles.

The interpretation of scientific traditions, and their formal codification, if that is to be accomplished, are essentially a task for scientists. As this discussion demonstrates, the problems of interpretation are manifold and if they are not solved they may severely hinder the progress of science. The harm done may not only be general but may apply particularly to industrial research of the group type. Incidental effects of the code, but of considerable importance, would be the great improvement in morale among scientific workers, the improvement in the quality of scientific work, the assistance it would give to editors of scientific journals and research administrators, and the basis it would provide for exposing poor work and even instances of chicanery. It would be of great

assistance in the training of graduate students in the scientific method. The preparation of a formal code of professional ethics should be of considerable value in establishing the professional status of the scientist in the public mind. It seems more than a coincidence that the groups that already have formal statements of their social and professional responsibilities and have definite rules of professional behavior are those definitely accepted by the public as having professional status.

As we have pointed out, violations of professional ethics on the part of scientists are frequent and familiar to all scientists. Sometimes they are deliberate violations for personal power or gain. Frequently, they are the results of carelessness or unfamiliarity of research administrators or research workers with the established traditions. They may even result from excessive pressure of work, a condition that appears common in industrial research. Some violations are the result of misguided attempts by editors and reviewers of scientific journals to shorten articles.

Is not the time opportune for our scientific organizations, or some agency of Unesco, to consider the manner of the application of scientific traditions to the newly developed conditions of scientific research? We suggest that the establishment of a definite code of professional ethics and conduct by our major scientific groups would have profound and favorable effects for science, society, and the scientist.

A mere statement of principles would be of help. An extensive codification and attempt to discipline or expose gross violations might be desirable. Our societies have various ways and means of enforcing regulations. Exclusion from membership and control of publications and means of publicity are powers that could be used to control unscrupulous and continuous violations. There may appear to be an anomaly in scientists' establishing a formal code of ethics to preserve traditions that include independence in their work, but this merely reflects an anomaly in present conditions of scientific work. It seems far better for scientists to affirm such a code positively than to be regimented to an increasing extent without any control over the conditions under which they must work. A. V. Hill (*8*) puts the problem as follows:

> The important thing is not a creed "which except a man believe faithfully he cannot be saved." What matters is that scientific men should argue and discuss the matter of scientific ethics as one of infinite importance to themselves and the rest of mankind with the same honesty, humility and resolute regard for the facts they show in their scientific work.
>
> If they do, then something will surely crystallize out from their discussion, and I have faith enough in the goodness and wisdom of most scientific men to believe that the result on the whole will be good and wise. It may in the end be embodied in a new Hippocratic Oath; or it may be absorbed in trade union rules for the scientific profession; or ethical behavior in science may just come to be accepted as an honorable obligation as unbreakable as that of accuracy and integrity.

We add that all problems will not be solved, but science is expanding and moving. The rate of progress will be profoundly affected by the consideration that is given to the maintenance and proper application of time-proved scientific traditions as the conditions of scientific work change.

References

1. Bernard, Claude. *An introduction to the study of experimental medicine.* Trans. by H. C. Greene. New York: Macmillan Co., 1927.
2. *Ibid.*, p. 14.
3. Blakeslee, A. F. *Science*, 1940, **92**, 589.
4. Carlson, A. J. *Science*, 1946, **103**, 377.
5. Condon, E. U. *Science*, 1946, **103**, 415.
6. Fernelius, W. C. *Chem. Eng. News*, 1946, **24**, 1664.
7. Gehrke, W. H., Anebhansley, C. H., and Rothemund, P. *Chem. Eng. News*, 1947, **25**, 2562.
8. Hill, Archibald V. *Chem. Eng. News*, 1946, **24**, 1343.
9. Leake, C. D. (Ed.) *Percival's medical ethics.* Baltimore: Williams & Wilkins, 1927.
10. Mills, C. A. *Science*, 1948, **107**, 127.
11. Peoples, S. A. and Carmichael, E. B. *Science*, 1945, **102**, 131.
12. Polanyi, M. *Sci. Mon.*, 1945, **60**, 141.
13. Wise, L. E. *Paper Industry & Paper World*, 1946, **28**, 1299.
14. *A.M.A. Directory*, 1942, p. 14.
15. *A.D.A. Directory*, 1947, p. 2.
16. *Bull. A. A. U. P.*, 1949, **35**, 66.
17. *A. I. C. E. Constitution*, Art. VIII.
18. Engineers' Council for Professional Development, Oct. 25, 1947.
19. *J. Hered.*, Sept. 1939.

Fraud in Medical Research

PRESIDENTIAL ADDRESS DELIVERED BEFORE THE 74TH ANNUAL MEETING OF THE AMERICAN SOCIETY FOR CLINICAL INVESTIGATION, WASHINGTON, DC, 8 MAY 1982

PHILIP W. MAJERUS, *Division of Hematology-Oncology, Washington University School of Medicine, St. Louis, Missouri 63110*

In the fall of 1963 my four-year-old daughter Suzie, who has since graduated from MIT, predicted my future. A neighbor had inquired over the back fence "What does your daddy do?" Suzie replied, "He used to be a doctor but he doesn't work any more." We had moved to Bethesda from Boston a few months earlier, and my job had changed from Resident in Internal Medicine at Massachusetts General Hospital to Research Associate in the Laboratory of Biochemistry at the National Heart Institute. Since that time I have not "worked," rather I have enjoyed the great fortune of doing what I please—medical research. I feel very fortunate to have been allowed to pursue problems on my own terms and am extremely grateful for our system of support for medical investigation. It is therefore with regret that today I discuss a vexing problem, that of fraud in medical research. I am revolted by anyone who fabricates results in the laboratory. I cannot explain such an individual nor can I justify his actions in any way. However, in approaching the question of fraud in medical research, I will attempt to put my personal revulsion aside. It is easy to overreact and it concerns me that we in the medical research community may do just that, even though I can understand why. It is hard to be dispassionate about something that strikes at the heart of our lives. However, by overreacting we may invite the formation of regulatory groups that will impede, rather than facilitate honest medical research.

In the past few years, there have been a half-dozen scandals involving outright fabrication of research data. These episodes have received wide press coverage, both because of intrinsic prurient interest in such news and because cheating in the use of a method where the goal is the discovery of truth is so aberrant. It's like trying to construct a building using dynamite—it does not work.

Why does the scientist cheat? The number of episodes of dishonesty seems to be increasing. Several factors may account for this. First, the number of investigators involved in medical research as their major career effort has increased enormously. At the second annual meeting of the American Society of Hematology held in St. Louis in 1959, there were 41 submitted abstracts; in 1981 there were 875. The membership of the American Federation for Clinical Research has tripled since 1962, and since 1968 the number of investigators in clinical departments of medical schools has increased from 11,000 to 30,000.

Of possibly more significance is the fact that there is a qualitative difference between today's investigators and those of the past. The physician-investigator used to be just that—a physician first and an investigator second. The stake in "results" was less. One had medical practice to fall back on if things didn't work out in the laboratory. This allowed for a gentlemanly practice of science which, although competitive, did not produce the pressures of current times.

As biology has become more technically complex, it has become more and more difficult to be both a physician and an investigator working at the forefront of medical research. Thus, to most of us, medical practice as a fallback position has become an increasingly unrealistic alternative to an unsuccessful investigative career. The pressure to produce is inevitable and although this does not *cause* cheating, it may contribute to a loss of perspective. We have all had the experience of the spurious spectacular result that leads to temporary euphoria, only to have our hopes dashed when the controls were completed or the artifact discovered. How tempting it is not to look too hard for that "fly in the ointment."

A second reason for the increased recognition of cheating scandals has come because the public's per-

ception of medical research has changed. When I entered the field, investigators were judged by the public as absentminded dreamers of little interest to men of affairs. If we did something of benefit to mankind, great! If not, no one really cared. Recent successes in clinical investigation and in molecular biology have convinced the public that biological research is useful and important. Thus, a new reward is potentially available to the investigator. He can become a "star" in the eyes of the public. No longer need he fumblingly describe what he does at family reunions to incredulous relatives, the epilogue being delivered by the patriarch: "Don't worry, he'll come to his senses and hang out his shingle soon."

The increased awareness in the activities of medical investigators means that our foibles are front page news and not just gossip for the laboratory coffee-klatch. I believe that striving for "stardom" is a perverse motivation for a scientist and it is a major factor in some cases of fraud. A number of other conditions contribute to the likelihood of dishonesty. These include the unscrupulous mentor who pressures fellows and students for "results" to obtain or maintain stardom, thereby suborning them, and the "absent" mentor who assembles large groups of students and fellows and sets them to work without adequate supervision, thereby failing to expose them to the skepticism and rigor important to the scientific method. A recent report on dishonesty in research by an *ad hoc* committee to the dean of the Harvard Medical School listed additional factors that could foster dishonesty, such as excessive publication of fragmentary results and multiple abstract submissions.

What is the significance of fraud to progress in medical research? Fraud affects medical research in two ways. The first is to sully the "white knight" image that the public has afforded investigators. There can be no doubt that recent scandals have tarnished our public image, albeit deservedly so. However, the public's increased interest in, and knowledge of, medical research means that they are beginning to know us better; thus, they will inevitably discover that we are no better or worse than other people. Therefore, I conclude that our loss of revered status, though regrettable, is inevitable. Not only is it inevitable, but, in fact, may be good. A public imbued with a healthy skepticism is a valuable ally.

The second, and potentially more important, way that fraud may affect medical research is to retard progress. Fabrication of research results has been of two types. In the first, the claimed results are of great significance to a fast-moving field, or of great theoretical or practical importance that would open up new areas for investigation. Thus far, in these instances where the damage would be potentially great, the false results have been refuted or the fraud discovered within a few months. In other cases, the work itself has been of little apparent significance being either primarily confirmatory or data-gathering that I categorize as "brownian motion research." That is, movement without progress. In such cases, one could argue that no one cares whether the work is right or wrong. Even the claim that money is wasted in refuting such frauds is discounted by my contention that investigators who are likely to make lasting contributions are unlikely to waste time on such projects.

In thinking of ways that fraud could hinder progress in medical research, I believe that the most damaging fraud might be in a field where the claims are significant, but the experiments are so complex as to take years to repeat. This is possible where the results of a particular therapy would take years to evaluate. Although at present I know of no such cases, their occurrence would be very serious. Probably the complexity of such experiments involving many investigators renders fraud unlikely. Dishonesty in such cases would require the conspiracy of many workers.

Fraud is a fabrication of results, either from "whole cloth", or by the techniques of "data management," such as, "buffing up the curves" or throwing out results that do not conform. However, a much more insidious danger faces the investigator than outright fraud, the problem of self-delusion. I mentioned earlier the spurious spectacular result that leads to temporary euphoria and the temptation not to look for the fly in the ointment. Human nature cannot be denied. We have all indulged in self-delusion at one time or another. However, it has no place in the laboratory. Self-delusion involves misinterpreting one's own results. This can follow from events as simple as omitting proper controls or failing to repeat complex experiments that conform to preconceptions. In fact, I believe that self-delusion occurs most frequently when an investigator anticipates the result of an experiment before it is performed. In this case, "aberrant" features of the results are either ignored or discarded as unimportant, when in fact they are the key to progress. I will give a hypothetical example to illustrate this point.

An experiment is designed to evaluate the effect of vitamin K on the in vitro synthesis of prothrombin by cultured liver cells. Cells are incubated with radioactive amino acids with and without vitamin K, and newly synthesized radioactive prothrombin is isolated by immunological techniques. Fig. 1 shows the results of four experiments with stimulation by vitamin K plotted vs. time. Note that in all four experiments a "positive result" is obtained. Also note that the synthesis of prothrombin with time is not really linear. In each experiment there is a "glitch" in the curve

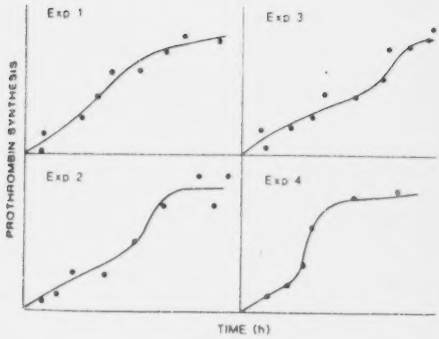


FIGURE 1

where it seems that prothrombin accumulates in a greater than linear manner at a variable time after institution of the culture. Our hypothetical investigator ignores this aberration and instead "normalizes" the data and averages the four experiments as shown in Fig. 2. A line is drawn through the points; the original hypothesis that vitamin K is required for prothrombin synthesis is validated. The work is then written up and a paper is submitted for publication. In the interests of "conserving space in the journal" the editors insist that only the summary graph be published. Thus, we have "lost" information to all but those evaluating the original data. Suppose another worker has the idea that degradation products of prothrombin that result from its proteolytic turnover serve to regulate prothrombin synthesis by stimulating its production. This hypothesis would predict that as prothrombin is synthesized and then degraded by the inevitable proteases present in tissue culture media that synthesis would accelerate. The second investigator might be dissuaded from pursuing the hypothesis upon examination of the published results of the first worker. Yet, in fact, the raw data of those experiments are consistent with the hypothesis. Thus, although no fraud has occurred, progress has been potentially retarded. In my five years as an editor of The Journal of Clinical Investigation I handled over 6,000 papers. I found dozens of examples of self-delusion and only two cases involving fraud.

What should be done to prevent dishonesty in science? The current wave of publicity about dishonesty in science has led the NIH, the American Association of Medical Colleges, and several institutions to form committees to study the matter. Questions addressed include: What can be done to prevent erosion of public confidence in the honesty of the biomedical research community? How can institutions promote ethical conduct in research? At what levels should measures be applied to prevent fraud—government? journals? institutions? individual laboratories?

In studying the recent cases I would argue that the current methods for dealing with dishonesty in science, although imperfect, have meted out stern punishment for wrongdoing. The guilty parties have been effectively excluded from further pursuit of careers in medical research, and even co-workers who may have been partially culpable to totally innocent have suffered irreparable damage to their reputations. I recommend that some additional policies should be adhered to in investigating cases of alleged dishonesty in science. These include: (*a*) retraction of all papers relating to the work in question, even if parts of those papers are not in dispute; (*b*) no further publication of any papers involving work of the accused until all issues have been resolved; (*c*) investigation of the details of the reported fraud by a group that does not include anyone from the involved laboratory or their close associates; and (*d*) suspension of any *new* grant awards to the accused until the issue is resolved. I do not want to discuss today how or by whom these policies should be applied. Rather, I will speak about methods for prevention of fraud. Can government work to prevent dishonesty?—Unlikely. The NIH bureaucracy is too far removed from the daily activities of the laboratory. Furthermore, experience with NIH remedies shows that the treatment is often worse than the disease. A recent example involves the guidelines for recombinant DNA research. Although the impetus for the guidelines came from molecular biologists themselves, in practice the guidelines almost hamstrung the most promising new avenue of scientific research in decades.

In that case, the *real* issues were technical. What are the likely benefits to society of recombinant DNA research vs. the likely dangers of the new technology? In retrospect, it is easy to see that the benefits were obvious and concrete while the risks were entirely hypothetical. Yet the controversy was quickly changed

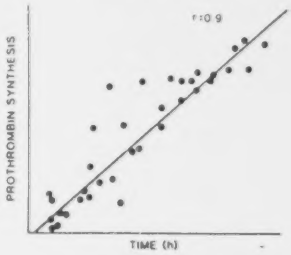


FIGURE 2

from the technical issues to moral and philosophical concerns. Various splinter groups sought to ban recombinant DNA research, not on the grounds of specific dangers but rather because of fears inspired by *Brave New World*. As James Watson said, "A tiny group of Boston-based academic leftists fantasize that the rich will finally subjugate the masses by giving them bad genes manufactured by recombinant DNA methodology."[1] Only through many hours of lobbying and education of politicians and administrators was it possible to continue the research and eventually to disarm the guidelines.

A lesson can be learned from this episode. A scientist who becomes inactive may lose his ability to judge technical issues. Since ex-scientists and other administrators could not validly judge the technical issues, they attempted to "solve" the DNA problem by guidelines. The guidelines were established and enforced by committees.

In my opinion, rigid rules or "guidelines" are often used as methods to avoid judging issues on their specific merits. This strategy is used by institutions wishing to avoid individual responsibility and by administrators who, as "supernumeraries" less capable than those they wish to control, use rigid rules to avoid confronting the substance of problems. Collective decisions mean that individuals are not directly accountable. We have all been frustrated by verdicts rendered by such groups as "the council" or "the editors." It reminds me of the firing squad where no one executed the victim because one gun had a blank. Therefore, I would argue that government has little to offer to the solution to the problem of prevention of dishonesty in research. The same arguments apply to a lesser degree to institutions, as a whole, such as universities or large research institutes.

Suggestions that journals take a greater role in ensuring against publication of false data are also impractical. We are the JCI not the FBI. The most that a journal can do if the reviewers or editors suspect that data are either false or, as is more often the case, wrong for less sinister reasons is to refuse to publish the work. Journals have no mechanism to investigate whether data are false or not. Even when fraud is discovered, journals are in a poor position to respond. In most cases of fraudulant papers, there are several authors. One or more is accused of cheating and the others are presumably innocent. Journals have no way of sorting out the guilty from the innocent. In my opinion when any part of a study is found to be fraudulant, the entire work should be retracted pending its repetition. This has not always happened in practice as authors have asked to retract only parts of the data in papers. In only one instance of a dispute between authors in my time with the JCI was I able to communicate with all of the authors of a disputed work, thereby sorting out the facts of the case. In that instance, fraud was not an issue. Let me illustrate the problem for journals by another hypothetical example. Suppose a paper with three authors is published in which a variety of types of measurements are made. Later, two of the authors write to the journal stating that the work of the third author is false and therefore should be retracted while the remainder of the work should stand. We hear nothing from the third author whose guilt or innocence has not been established. If the journal publishes a conditional retraction of the work, the third author is convicted of wrongdoing without any hearing. As recent scandals have illustrated, such a conviction will destroy the reputation of this worker. Thus, other mechanisms are required to prevent and punish fraud. The journals can play only a passive role in this regard.

What can be done to prevent publication of false work? I believe that the only level at which any effective measures can be applied is in the individual laboratory. Students, postdoctoral fellows, and junior colleagues are the future of medical research. They are our most valuable resource and should be treated as such. Senior investigators have a solemn responsibility to guide trainees to allow them to express their full potential. If, because of clinical, administrative, or other constraints, an investigator does not have the time to participate in the ongoing progress of an investigation on a *day-to-day basis*, then he should dissociate himself from it. It is another type of self-delusion for a senior investigator to claim "It was my idea" or "I wrote the paper" or "It's my grant" as justification for taking trainees into the laboratory but not directly participating in the work. The contribution of the senior investigator in such cases is miniscule. It's rather like claiming credit for writing *Hamlet* because you furnished Shakespeare with a pencil. In such laboratories the training environment may still be good if advanced trainees or technicians are outstanding, but that is uncommon in my experience. The worst aspect of this situation is that lack of supervision leads not to dishonesty but more often to lack of productivity and progress due to the investigation of trivial problems.

Work in progress should be discussed openly and the data should be reviewed frequently, not just by the laboratory chief but also by disinterested parties. Group meetings of large laboratories where there is evaluation of data of individuals are important. Even better are presentations to departmental or other groups where investigators not directly connected with the work evaluate the data. These exercises require

[1] James D. Watson and John Tooze, *The DNA Story* (San Francisco: W. H. Freeman, 1981), p. 383.

heavy applications of skepticism, the most important ingredient in scientific creativity. As Gerald Holton has pointed out, progress in science is generated by individuals who refuse to accept dogma and thereby develop new constructions that open avenues of research.[2] One must be highly tuned to receive signals from aberrant results.

One of the most repeated excuses given for dishonesty in the research laboratory is the theory of "publish or perish." The end result of the publish or perish syndrome is much more insidious than the occurrence of a handful of fraudulent laboratory incidents. Each year talented young people are lost to clinical investigation because the number of papers published takes precedence over the quality and intensity of the effort expended.

Progress in science is episodic and discontinuous. Thus, over the long haul a creative and persistent investigator is certain to produce new information. However, the pace of discovery depends on vagarious factors. Suppose two equally talented individuals are working on hypotheses, one of which is correct, the other incorrect. There is no *a priori* reason to conclude that one hypothesis is more likely to be correct. In this case, the person with the correct hypothesis will get "results," the other will fail. Thus, in the short run one person appears better than the other, in the long run it will even out. We are in a marathon, not a 100-yard dash! Young investigators must be supported long enough to reach their level of achievement without undue pressure to produce. Since most administrators can only count publications, there needs to be someone in the chain of command who judges the investigator on more substantive grounds, and his opinion should override that of the administrator.

Until now the occurrence of fraud has damaged medical researchers primarily by tarnishing their image, but it has not retarded progress in any substantial way. I believe that efforts to curtail the publication of incorrect results can only effectively arise at the level of individual laboratories or small groups. You will notice that one theme keeps recurring—that of close and active participation in the laboratory by the senior researcher. Dishonesty in research is antithetical and therefore inexplicable, but it can be curtailed. The assembly line mentality and production of results for the sake of publishing only have no place in medical research. The intrusion of governmental and institutional regulatory committees will only retard progress, and the creation of paper guidelines will serve no purpose. The scientific method is self-correcting and no drastic remedies are required. As Pogo said, "We have met the enemy and they are us."

[2] Gerald Holton, *Thematic Origins of Scientific Thought* (Cambridge: Harvard University Press, 1973).

Authorship from the Reader's Side

THE WORD *author* comes from a Latin verb meaning to "increase, promote, or originate" (1). Indeed any reader should expect to find in a scholarly article something added to the sum of knowledge. But in recent years, the "increase" in each medical paper is beginning to look less like an increase in knowledge and more like an increase in the number of authors. Evidence for such a trend is pointed out in this issue by Burman in his "Perspective" article on authorship (2), and he discusses reasons for this apparent increase.

The growing complexity of medical research is responsible in part for an increase in the number of authors per paper. Another possible cause is a growing diversity of authors. We are seeing more papers with authors having academic degrees other than the M.D.; these papers with diverse authorship do not appear to have a lower level of quality. This apparent trend suggests that the number of potential authors is rising in diverse disciplines, with growing competition among them and with physicians for authorship in biomedical journals.

The main question raised by Burman is not numbers of legitimate authors for papers but the legitimacies of authorship and place among authors for each author listed on a paper regardless of the number. Unquestionably, authorship is important, economically important. Burman's use of the word "masthead" in his title has led me to think of papers in preparation as academic life-rafts onto which potential authors scramble for a place, another chance to survive in the stormy sea of academic competition. Burman's analysis does not consider in detail the question of who is entitled to a place on the raft, but concerns itself more with who gets what place on the raft once a place is assured.

Others are also calling for definition of authorship (3). Eugene Garfield, publisher of *Science Citation Index*, has recently raised his concerns in an address to the Council of Biology Editors (4). An ad hoc committee of the Association of American Medical Colleges concerned with fraud in medical research has asked, in its recently issued guidelines (5), that institutions examine their "policies on authorship of papers and abstracts to ensure that named authors have had a genuine role in the research and accept responsibility for the quality of work being reported."

Who is going to define authorship in terms that will satisfy not only claimants to authorship but readers and editors as well? Could authors do it? What body of authors? Is any one scholarly society in medicine representative enough of all medical authors to formulate a consensual definition that will be seen as ethically binding by all potential authors? An organization like the American Federation for Clinical Research might take up the challenge but would its influence be wide enough? Garfield told the Council of Biology Editors that editors "have the authority to establish guidelines for authorship." I think Garfield is right, not because of editors' authority but their sensitivity. Guidelines responsive to the interests of all parties would take into account the needs and expectations of both authors and readers, balanced fairly to the extent that the needs might clash. This sensitive balancing of interests is already a central duty for editors. We try to make a place in our journals for authors and help them present their articles effectively. For readers we try to select papers that will serve their various needs and take steps, such as peer review, to try to supply them with useful information presented honestly. How, then, might editors define authorship to give contending potential authors guidelines with which they could decide among themselves who should be authors while satisfying the needs of readers?

I suggest we start with the readers. I have commented elsewhere that when readers look at a paper they want to know who is taking *public responsibility for its content* (6). Three words in this definition can be expanded into more detailed guidelines. *Responsibility* means the ability and willingness to defend the content of the paper if it is challenged by readers. *Public* means that authors are willing to carry out this responsibility in a published defense, such as a signed letter to the editor; private defense in private correspondence would not reach the scientific public. *Content* means not simply packages of data but also the conceptual framework on which they are hung: the justification for a study or clinical observations; the basis for the study design; methods for collection of valid data; the analysis and interpretation of the data; and the logic that led to the conclusions.

From these points follow specific rules that potential authors should be able to apply to themselves or their contending colleagues:

1. An author should have participated in the initiating or planning of a study or have assented to its design if enlisted late in the study.

2. An author should have made some of the reported observations or generated some of the data.

3. An author should have participated in interpreting the observations or data and deriving from them the reported conclusions.

4. An author should have taken part in the writing of the paper.

5. An author should have read the entire contents of a paper and assented to its publication before it is sent to a journal.

One need not ponder long over these criteria to think of legitimate exceptions. There is, for example, the statistician who designed a study of a drug intervention for a hyperlipoproteinemia, monitored the collection of the data, and applied elegant statistical methods to analysis and interpretation of the data. The statistician does not understand in depth the relevant biochemistry and could not explain why particular chemical methods were used in collecting the data. Yet he or she could defend some critical aspects of the study to the scientific public, as well as its design and some interpretations of the data.

If these criteria follow logically from the premise that an author is a person who can take public responsibility for the content of a paper, and this premise is accepted,

then various exclusions from authorship are proper. The radiologist who happened to take and interpret routine roentgenographs mentioned in a case report would not be a legitimate author unless those interpretations led to new concepts. The laboratory technician who did the routine tests that yielded data interpreted for new concepts by others is not a legitimate author. The head of a department who did nothing more than approve a final version of the paper is not a legitimate author. Consider that the person entitled to put his name on a mosaic is not he who gathered the chips but the artist who put together the picture; the two persons may be the same but the claim to creation lies in the picture.

Most of my discussion here does not hinge on new points. The central criterion for authorship was finely stated some years ago by the head of the Section of Publications at the Mayo Clinic, Richard Hewitt (also cited by Burman) (7):

> Authorship cannot be conferred; it may be undertaken by one who will shoulder the responsibility that goes with it. To a responsible writer, an article, with his name on it, is the highest product of his mind and art, his property, as nearly flawless as he can make it, founded in his character and evidence of it. If that describes the acceptable standard, medical writers, a responsible group, are in present need of reconsidering the implications of joint authorship. The reader of a report issued by two or more authors has a right to assume that each author has some authoritative knowledge of the subject, that each contributed to the investigation, and that each labored on the report to the extent of weighing every word and quantity in it.

This may sound a bit old-fashioned. But even if Hewitt was not writing in the context of today's scholarship and research, is he asking too much for today's authors?—THE EDITOR.

References

1. ONIONS CT, ed. *The Oxford Dictionary of English Etymology*. Oxford: Oxford University Press; 1966: 63.
2. BURMAN K. Hanging from the masthead: reflections on authorship. *Ann Intern Med*. 1982;97:602-5.
3. BROAD WJ. Crisis in publishing: credit or credibility. *BioScience*. 1982;32:645-7.
4. GARFIELD E. More on the ethics of scientific publication: abuses of authorship attribution and citation amnesia undermine the reward system of science. *Curr Contents*. 1982;25(30):5-10.
5. AD HOC COMMITTEE ON THE MAINTENANCE OF HIGH ETHICAL STANDARDS IN THE CONDUCT OF RESEARCH. *The Maintenance of High Ethical Standards in the Conduct of Research*. Washington D.C.: Association of American Medical Colleges; 1982.
6. HUTH E. Preparing to write. In: *How to Write and Publish Papers in the Medical Sciences*. Philadelphia: ISI Press; 1982: Ch. 3.
7. HEWITT RM. *The Physician-Writer's Book: Tricks of the Trade of Medical Writing*. Philadelphia: W. B. Saunders Company; 1957:312.

PERSPECTIVE

"Hanging from the Masthead": Reflections on Authorship

Lt. Col. KENNETH D. BURMAN, MC, USA: Washington, D.C.; and Bethesda, Maryland

Authorship of a scientific article should imply expertise on its content and thorough knowledge of the investigation reported. Because the number of authors per article has dramatically increased, I question whether these criteria for authorship are applied. Another explanation for this increase is that authors add names to the byline without reference to any set of criteria, resulting in the designation of coauthorship when not warranted. The medical community needs to redefine the proper procedure and criteria for deciding on authorship and to strictly abide by these procedures. Specific recommendations center around the protocol as the initial instrument of communication. Who should be designated as author, and the order of names, should be negotiated before and during the study, and written communication with other involved laboratories should result in more appropriate authorship attribution.

> The publication of a long list of authors' names after the title is a little like having all a vessel's ballast hanging from the masthead, as if to counterbalance the barnacles(1).

AUTHORSHIP of an article, in the past, has implied, at least to one expert on medical writing (2), a substantial knowledge of its subject, participation in the investigation it presents, and thorough familiarity with its content. Are these guidelines still being adequately applied to define who should be designated as an author? An apparent trend toward a higher number of authors per article could be accounted for by inclusion of persons as coauthors who had not contributed substantially enough to merit this designation.

We are all aware of the increasing number of authors per article in medical journals. A detailed study of this phenomenon showed that the average number of authors per article in all scientific fields (based on the *Science Citation Index*) climbed from 2.01 in 1970 to 2.31 in 1975 and increased to 2.50 in 1979 (3). This tendency is especially obvious in the field of life sciences. King and associates (4) found that in 1979 there were 2.01 authors per article on the life sciences whereas most other fields had lower figures; articles on the social sciences, for example, had an average of 1.35 authors, and the average for several fields was 1.79 authors per article. In an examination of the frequency of multiple authorship (5), 94.7% of all articles submitted to the journal *Cancer Research* had multiple authors and 5.3% had one author; 66.8% of all articles submitted to the *Astrophysical Journal* had multiple authors and 33.2% had one author. The average number of articles published per scientist in the life sciences field was 0.3 in 1965 and was estimated at 0.42 in 1980; these values contrast, respectively, with the figures of 0.6 and 0.22 articles per scientist in the social sciences field, and with 0.02 and 0.03 articles for those in the field of engineering (4). This trend of increasing numbers of authors per article, especially in the life sciences field, does not seem to correlate with an increasing frequency of citation, one measure of acceptance by the scientific community. A recent study in chemistry indicates an increasing number of authors did not enhance the likelihood of citation of that article. For example, of the 50 articles cited most frequently in the year 1972, 18 had only one author, only four papers had as many as four authors, and no paper had more than four authors (6). This lack of correlation occurs in several other fields (7), as the median citation counts (the number of times a particular article is cited in a reference list) in the field of economics from 1970 to 1975 for articles with one, two, and three authors were 1.2, 3.3, and 0.5, respectively (7). Similarly, the median citation counts for articles with one, two, three, and four or more authors were 1, 0.7, 4 and 3.2 in psychiatry; and 4, 8, 12.5, and 4.5 in biochemistry, respectively (7). Although these data were not corrected for the actual number of articles published with a designated number of authors, there still does not appear to be a relation between the numbers of authors and the frequency with which that article is cited.

Multiple Authorship in Medicine

An analysis of trends in authorship in *The New England Journal of Medicine* and *Annals of Internal Medicine* is shown in Table 1. The greatest rise in authors per article has occurred in the last decade both for case reports and original articles, and the numbers of authors per article in both journals were approximately twice as high as the average obtained for 1979 from *Science Citation Index* for all fields. A reader, of course, usually cannot discern the reasons why a particular medical article might have numerous authors. One example is the published report of a study that investigated the effect in five patients of the infusion of plasma that had been treated by being perfused over immobilized protein A(8); this report has 17 authors. Similarly, what might have been the contributions of each of the 10 authors of the case report of one patient who was noted to have a monoclonal antibody to peripheral nerve myelin(9)? In some circumstances, the basis of multi-authorship of articles in clinical medicine is clear. For example, clinical research generally requires a patient. The physicians who made the diagnosis and had been caring for the patient in one medical center may have to refer him or her to another medical center. This process might then require that the contribution of the referring physicians be acknowledged in the form of coauthorship. Also, because of the degree

From the Kyle Metabolic Unit, Department of Clinical Investigation, Walter Reed Army Medical Center, Washington, D.C., and the Department of Medicine, The Uniformed Services University of the Health Sciences, Bethesda, Maryland.

Table 1. Number of Authors (Mean ± SE) per Article in *Annals of Internal Medicine* and *The New England Journal of Medicine*[*]

Year	*Annals of Internal Medicine*†		*The New England Journal of Medicine*†	
	Original Article	Case Report	Original Article	Case Report
1930	1.3 ± 0.1(27)	...	1.2 ± 0.1(25)	...
1969	3.2 ± 0.2‖(54)	2.8 ± 0.2(36)	3.8 ± 0.3‖(25)	2.7 ± 0.2(27)
1979	4.7 ± 0.4§(38)	4.0 ± 0.2‡(62)	5.2 ± 0.4§(28)	3.2 ± 0.4(20)

* This compilation does not include special articles, reviews, or obvious multicenter cooperative trials. Volumes 4, 70, and 90 were analyzed for *Annals of Internal Medicine* and volumes 203, 280, and 301 for *The New England Journal of Medicine*. When the data are expressed as the increase in the number of authors per article per year, the results for *Annals of Internal Medicine* are 0.0487 (1930-1969) and 0.150 (1969-1979); for *The New England Journal of Medicine*, 0.0667 (1930-1969) and 0.140 (1969-1979).

† The numbers of articles surveyed are shown in parentheses.

‡ $p < 0.001$, compared with the respective value in 1969.

‖ $p < 0.005$, compared with the respective value in 1969.

§ $p < 0.005$, compared with the respective value in 1930.

of medical subspecialization, various sophisticated laboratory determinations may be done in several laboratories, which then must cooperate to collate these results. These factors might serve to increase the mean authorship number in contrast to that in biochemistry, where, one person might conceive, execute, and write the results of a given study, and where the opportunity and need for collaboration may be less.

Nevertheless, many readers in clinical medicine may be wondering whether this well documented increase in authors per article also represents an increasing inclination to designate persons as coauthors when, in fact, their contribution was not sufficient to warrant this designation. Very likely the explanations why so many reports have many and, perhaps, excessive authors are complex. Additional factors such as egocentrism and self-aggrandizement cannot be excluded. It is probably widely believed that inclusion of a person's name as author will allow him or her to expand the bibliographic citations on his or her curriculum vitae and, presumably, on a long-term basis, allow certain privileges, such as easier access to new jobs, promotions, and election to prestigious medical societies. The universities and societies that bestow these honors should not be criticized for relying so heavily on the content of curriculum vitae; a curriculum vitae, when properly conceived, does document to a degree a person's initiative, experience, and expertise in certain areas. Yet when that curriculum vitae represents credit (that is, articles) that are inappropriately assigned to that person, misleading conclusions on that person's achievements and originality are derived, and in some cases, the perquisites given to that person are inappropriate.

Selection of Authors for Multi-Authored Papers

As a general guideline, Dr. Morris Fishbein (10) advocated that "if a person has rendered service of major importance, his name should be included with those of the authors of the paper." This advice seems appropriate but leaves unanswered the basic issue of who should decide if the criteria for inclusion as a coauthor have been met. Further, how are such issues resolved when people disagree on both the general criteria for authorship well as the extent of specific contribution by a given person. Could a person who actually did most of the work, for example, a fellow in training, inform his senior that his name should not be included on an article? Indeed, could a fellow in training, perhaps participating in his first clinical investigation, actually decide by himself who is entitled to authorship? On the other hand, a "senior author" who may be the head of a laboratory might be so remote from a given project's daily progress that it may not be appropriate that he decide authorship. Indeed, the senior author's name might not even deserve to be included; it is an open secret that some chiefs of laboratory units require that their names be on all manuscripts sent in for publication, regardless of the magnitude of their input into a given project.

Negotiations on inclusion of a person's name as an author of a manuscript, however difficult, often are easier than the negotiations on the position of the names in the authorship statement. Generally, the reader may assume correctly that the first author is the one who actually did most of the work in a clinical investigation study. For example, first authors of articles submitted to the *Journal of the National Cancer Institute* spent 53.5 hours writing the paper whereas coauthors spent an average of only 19.5 hours on this task (11). In my experience, however, another person such as the "junior" staff man or assistant professor has occasionally usurped this first-author position, especially when a fellow in training did the study. Needless to say, the fellow had little or no recourse in such an action.

Investigators themselves do not agree on which positions after the first one are more desirable. In my experience, some people who actually guide and intensively participate in a study prefer that their name be last, because they believe this is the "senior author" position and would be recognized as such. Others believe that the second position is preferable, especially because of the increasing tendency of journal references to only cite the first three authors and attribute the remainder of the study participation to "et al." There does appear to be agreement, however, among the scientific community that the contribution of all authors other than the first, second, and last was minimal, when, in fact for the reasons mentioned above, this conclusion may be erroneous.

Cooperative Studies

The difficulty in deciding authorship is compounded in

Table 2. Suggested Placement of Authors' Names Based on the Extent of Participation*

First Author
- Clinical Investigation
  1. Planned and did study, evaluated data, and wrote the paper.
  2. Did the study, evaluated the data, and wrote the paper.
- Case Report
  1. Cared for patient, recognized the unusual nature of the disease process, made the diagnosis, planned protocol of study, and wrote the paper.
  2. Patient referred by another physician but made diagnosis, planned protocol, and wrote the paper.
  3. Planned protocol and wrote the paper.
  4. Carried out the protocol and wrote the paper.

Last Author
- Clinical Investigation
  1. Planned and supervised the study and the writing of the paper.
  2. Supervised planning, performance, and writing.
  3. Supervised either planning, performance, or writing.
  4. Procured funding by writing original protocol.
- Case Report
  1. Planned study protocol and supervised its performance and the writing of the paper.
  2. Supervised study protocol, its performance, or writing of the paper.
  3. Recognized the unusual nature of the disease process and made the diagnosis.

Other Authors
- Clinical Investigation
  1. Participated substantially by either doing or supervising any stage of planning, execution, data analysis, or writing.
- Case Report
  1. Participated substantially in any stage of planning, studies, their execution, or data analysis or writing of the paper.

* These criteria are listed in decreasing order of importance of contributions.

considering the level of participation of colleagues at different institutions. Large-scale, multicenter, cooperative studies are frequently needed to accrue enough patients to study a given aspect of a disease. Analyses of these data require an immense amount of work. In many instances, only the name of an institute or group sponsoring such a study may be given as the author with all the participants names cited as a footnote on the title page. Some persons, however, such as the coordinator, may have lent a tremendous amount of time and effort to this endeavor, even greater perhaps than the sum of that expended by all of the other participants, and yet not receive appropriate credit.

Studies done as a result of cooperation between several persons or groups at different institutions can lead to misappropriation of credit and to arguments on authorship. There are many degrees of cooperation between groups ranging from simply measuring several samples to extensive involvement in planning and execution of the study, and data analysis. As a result, participation could result in coauthorship, a simple acknowledgment, or occasionally no recognition at all.

Occasionally an investigator is included as a coauthor against his or her wishes (12). This type of misattribution of collaboration could occur only as result of poor communication. There are various reasons for not wanting one's name included on an article, but these could include the honest belief that the participation was not sufficient to warrant inclusion, or that the results obtained and conclusions reached were not congruent with one's scientific objectives or philosophy.

Recommendations

Perhaps the medical community needs to define for itself in more detail than Dr. Hewitt (2) could in 1957 the proper procedure and criteria for deciding on authorship. I offer these recommendations as a step toward such definition.

1. A person's name should be included as a coauthor only if he or she contributed significantly to the scientific formulation or execution of a study, or to the writing of the paper reporting the study.

2. The individual responsibilities of the involved parties should be made clear at the initiation of a protocol and study. Because most, if not all, studies require a detailed proposal or protocol for either scientific or administrative reasons, this protocol should clearly state the role of the participants and should list the authors' names in the tentative order in which they are expected to be listed in the final publication. The importance of these issues is critical to the working relationship that develops and that subsequently leads to a successful (and harmonious) completion of a study. Discord resulting from unshared expectations can lead to the ruinous outcome of an otherwise well-planned study. Any changes in roles or responsibilities must be openly discussed as they develop, with appropriate alteration and realignment in the listing of authors' names.

3. Because there is no agreement on assignment of appropriate participation by the current order of authors, I suggest that the first author name should indicate the person who actually did and arranged most of the work and who also actively wrote and referenced most of the paper (Table 2). The second name should indicate the person who actually supervised the activities of the "first author" as well as helping with ordering supplies, planning the study, and writing the paper. In brief, the person who quantitatively did the second most amount of work should be listed second. The last name should indicate the person in whose laboratory the study was done and who was peripherally involved with the details of the study, but who also participated in either the general conception, supplying the administrative support, or overseeing the general progression of the study. All other authors' names should be listed after the second author in decreasing order of contribution to the study. Moreover, the first author should be the final arbiter of conflicts on the order of authors' names, their roles in the study, as well as where and when to send an abstract or a final paper. This role of the first author may occasionally conflict with that of the second or last author, and in such cases negotiation is needed. Obviously, the persons who actually received the funding for the laboratory or the specific proposal should have an important role in these discussions, and inclusion of their names as a coauthor would be warranted.

4. The term "responsible author" should be introduced

to refer to the person who supervised the planning and execution of the study and the writing of the paper. This person should also be responsible for insuring that there was adherence to the protocol, that the data are accurately presented, and for insuring that each author is informed on the status of the study. Correspondence from editors should be addressed to this responsible author, who in most instances would be the first author. The person who is to be designated the responsible author should clearly state this in the cover letter accompanying the manuscript.

5. Individual authors should have the right of the final decision on whether their names should be included as coauthors. The reasons a person might want his or her name deleted from the authorship statement vary; but such requests should be honored in all instances. An important corollary to this general suggestion is that all potential coauthors must receive a final draft of the paper before it is submitted for publication. Further, all data supplied to another laboratory and then reported in a paper must have final approval for publication by the laboratory that did the tests or assays. Authors in one institution may write up studies and misinterpret or misuse laboratory data, implying it has the sanction of the laboratory that did the tests or assays. A laboratory that supplies data should probably include a statement in a letter to the cooperating investigators that the laboratory must receive a copy of the paper before it is submitted for publication.

6. Editors of journals should require that the first author's transmittal letter contain a statement that all authors contributed to either planning, doing, or analyzing the study or writing the paper. Further, there should be no acknowledgments to persons without their written permission.

ACKNOWLEDGMENTS The author thanks Dr. Morton V. Malin, the Institute for Scientific Information, Philadelphia, Pennsylvania; Dr. Dennis D. McDonald, King Research, Rockville, Maryland; Ms. Helene Ebenfield, the National Science Foundation, Washington, D.C.; and Dr. Leonard Wartofsky, Timothy Boehm, and William Fears for their cooperation, advice, and counsel; and Mrs. Barbara Kuffler and Mrs. Estelle Coleman for secretarial assistance.

The opinions or assertions contained herein are the private views of the author and are not to be construed as official or reflecting the views of the Department of the Army or the Department of Defense.

▶ Requests for reprints should be addressed to Lt. Col. Kenneth Burman, MC USA, Ward 47, Clinical Investigation, Kyle Metabolic Unit, Walter Reed Army Medical Center, Washington, DC 20012.

References

1. Controversies in medical communication [Editorial]. *N Engl J Med.* 1964;271:1068.
2. HEWITT RM. *The Physician-Writer's Book: Tricks of the Trade of Medical Writing.* Philadelphia: W.B. Saunders Company; 1957:315.
3. *Science Citation Index.* Philadelphia: The Institute for Scientific Information.
4. KING DW, MCDONALD DD, RODERER NK. *The Journal in Scientific Communications: The Roles of Authors, Publishers, Librarians and Readers in a Vital System.* Rockville, Maryland: King Research; 1979; NTIS publication no. 296-263.
5. MCDONALD DD. *Final Report: Interactions between Scientists and the Journal Publishing Process.* Rockville, Maryland: King Research; 1979; DHEW publication no. (NIS) 293-990.
6. SMALL HG. *Characteristics of Frequently Cited Papers in Chemistry.* Philadelphia: The Institute for Scientific Information; 1974.
7. LINDSAY D. *The Scientific Publication System in Social Science.* San Francisco: Jossey-Bass, Inc.; 1978.
8. TERMAN DS, YOUNG JB, SHEARER WT, et al. Preliminary observations of the effects of breast adenocarcinoma of plasma perfused over immobilized protein A. *N Engl J Med.* 1981;305:1195-9.
9. LATOV N, SHERMAN WH, NEMNI R, et al. Plasma-cell dyscrasia and peripheral neuropathy with a monoclonal antibody to peripheral-nerve myelin. *N Engl J Med.* 1980;303:618-21.
10. FISHBEIN M. *Medical Writing: The Technique and the Art.* New York: McGraw-Hill Book Company; 1957:32.
11. KING DW, MCDONALD DD, OLSEN CH. *A Survey of Readers, Subscribers and Authors of the Journal of the National Cancer Institute: Final Report.* Rockville, Maryland: King Research; 1978.
12. BROAD WJ. The publishing game: getting more for less. *Science.* 1981; 211:1137-9.

Honor in Science

C. Ian Jackson
John W. Prados

Possible roles for Sigma Xi in regard to scientific fraud, irresponsible authorship, and "publish or perish"

Views

As Sigma Xi prepares to mark the centennial of its establishment at Cornell University in 1886, the Society is attempting to redefine its role as an honor society in a scientific research environment totally different from that in which it was created. For example, many members have argued that Sigma Xi should become more active in advocating increased government funding for scientific research. Others urge that the Society should give more direct support to research through its grants-in-aid program, or should use its structure of local chapters and clubs to influence the quality and status of science teaching at the high school level.

We shall not attempt to evaluate here the merits of such proposals, except to say that they deserve to be debated and that in fact the debate is going on. We do, however, propose another role for Sigma Xi that seems to us more vital and immediate than any of these, but that has attracted much less attention. Sigma Xi has always been an honor society of research scientists, and current evidence suggests to us that the threats to the integrity and honor of scientific research are at least as significant as they were in 1886. The most obvious and publicized of these threats are the instances of fraudulent research, including selective reporting of results, fabrication of data, and plagiarism in dissertations and publications. Because of the disproportionate effect that such frauds have on the overall credibility of scientific research, we believe that Sigma Xi cannot avoid giving this issue a high priority. However, if we begin to ask how fraud was possible and what the factors were that encouraged the attempts, we find answers that suggest other urgent tasks for our honor society. We propose that the Society undertake these tasks, because we are convinced that Sigma Xi's influence can be considerable.

Scientific fraud

Much recent discussion has been concerned with the amount of deliberate fraud in scientific research. Nicholas Wade, coauthor with William Broad of *Betrayers of the Truth* (*1*), has suggested that the discovered cases represent only the tip of a very large iceberg; this assertion was angrily disputed in a discussion at the American Association for the Advancement of Science meeting last May (*2*). Without attempting to estimate the size of the iceberg, we agree that scientific fraud is probably attempted much more frequently than it is detected, for three reasons. First, detection is most probable where the prospects of fame or fortune are very great, for example in cancer or other medical research. Because there are many research teams and strong public interest, there is a higher probability that replication of experiments will be attempted. If the perpetrator of a fraud works in a more obscure field and has more modest ambitions—university tenure, for example—the chances of detection are probably much less. Second, scientists are subject to the same temptations that afflict the general population. Although we may feel that scientists, as members of a trained profession, are particularly sensitive to ethical principles, such principles cannot guarantee behavior.

The third reason much scientific fraud may go undetected is that the mechanisms for preventing it have been shown to be more unreliable than scientists have traditionally wished to believe. Arnold Relman, editor of the *New England Journal of Medicine*, has recently expressed his increasing concern:

> What kind of protection against fraud does peer review offer? The Darsee affair gives a clear answer. Little or none. Most of Darsee's work was published in peer-reviewed journals, some with very exacting standards, yet in none of the reviews was there enough suspicion to warrant rejection. . . . Unless a maladroit cheat fabricates results that are manifestly impossible or inherently contradictory, even the most rigorous peer review is not likely to uncover fraud [*3*].

What could Sigma Xi do to deter fraud? We suggest the following.

1. The Society can make the integrity of scientific research a more visible priority than it has been in the past. Sigma Xi has taken honesty in research as something so vital that it has not needed mention. This may have been a mistake. Other professions have adopted explicit codes of ethics as guides to responsible professional behavior. Perhaps Sigma Xi, as a respected organization of research scientists, could propose a code of ethics for scientific research.
2. Sigma Xi could revise the pledge that it now requires of all those elected to membership. At present, members agree "to uphold the principles of Sigma Xi, The Scientific Research Society, to assume the responsibilities of membership herein and, so far as possible, to encourage original investigations in Science." It would not be difficult to make these principles and obligations more explicit, particularly as regards scientific honesty.
3. A mechanism could be created through which Sigma Xi membership can be revoked. At present, election to the Society is for life. Presumably revocation of membership would be rare, but we believe that an honor society of research scientists should be prepared to withdraw the honor of

C. Ian Jackson is Executive Director of Sigma Xi and Publisher of American Scientist. John W. Prados, President of Sigma Xi in 1983–84, is Vice-President for Academic Affairs and Research at the University of Tennessee and Vice-Chairman of the Engineering Accreditation Commission, Accreditation Board for Engineering and Technology, Inc. Address for Dr. Jackson: 345 Whitney Ave., New Haven, CT 06511.

63-471 - 87 - 18

membership in cases where the Society's principles or ethical standards have been ignored.

4. Individual members of the Society could be more rigorous in acting as referees of scientific papers, and in questioning doubtful data or results. "Whistle-blowing" is a difficult and unpleasant task; even when justified by the results, the whistle-blower may feel little satisfaction and may receive at least implicit criticism from some unthinking fellow scientists. Yet the credibility of science seems to depend on honest skepticism. Hans Christian Andersen does not relate whether any scientists were among those who made no comment on the emperor's new clothes, but it seems very possible. If individual members are subjected to criticism because they question scientific data on reasonable grounds, Sigma Xi should support them.

Since Sigma Xi is itself composed of fallible individuals, such measures can in no sense be regarded as a cure. But they could deter those tempted to fraud, and encourage those who do their research honorably. Implicit toleration of dishonesty in research threatens not merely the foundations of science but the welfare of society as a whole.

Irresponsible authorship

In their recent book, Broad and Wade suggest two principles to deter fraudulent research that concern "that critically important part of a scientific paper—the authorship line." They recommend that, "first, all people named as authors should have made a definably major contribution to the work reported. Any minor contribution should be explicitly acknowledged in the text of the article. Second, all authors of a paper should be prepared to take responsibility for its contents in precisely the same measure as they stand to take credit" (*4*).

We have two general comments on this suggestion. The principles evidently have much wider significance than merely as deterrents to fraud. Further, if these principles are no longer taken for granted in scientific publication, then the prestige of scientific research deserves to be attacked.

In his editorial on the Darsee affair, Relman restated these principles in similar language, and went on to point out some of their implications:

> Coauthors should be able and willing to defend the paper in public, and that means they must be confident about the integrity of the data. Furthermore, coauthorship should never be conferred or accepted as an honor or simply as a reward for providing resources or sponsoring a junior investigator who has done all the work. . . . Multiauthored papers should never be submitted for publication without the concurrence of all authors [*5*].

Although practically all the issues that we deal with here are encountered across the scientific alphabet from anthropology to zoology, it appears that the proliferation of authors of individual papers, and the reluctance of such individuals to accept responsibility for the entire paper, are most evident in biomedical research. There are factors related to the organization of research enterprises in these fields that partially explain this, but they do not justify such irresponsible authorship, which could extend rapidly and disastrously into other disciplines unless checked. Like many other scientists, we hear from time to time of situations in which one research scientist may name a colleague as a coauthor even though the latter has contributed little or nothing to the work reported, because the colleague will return the "favor" and both individuals' lists of publications expand accordingly. We also believe that some laboratory heads expect their names to appear on all papers published from the laboratories, even though this lacks significance in terms of scientific contribution or of scientific fame. The head of a successful research team is credited for the team's work by those whose opinions matter, without need for false authorship.

Some of the side effects of multiple authorship would be laughable if they were not pathetic. Take, for example, the matrix suggested in a letter to *Science* in 1981 (*6*), proposing an allocation of "points" to up to six authors of a research paper, according to the order in which they are listed (e.g., for five authors, the first-named is "awarded" 44 points out of 100; the fifth receives only 9). The questions that beg to be answered are "who counts the points and what is done with them?" but we dread the possible answers.

The worst effect of multiple authorship remains, however, the incentive it gives for irresponsibility. As Dr. Relman noted about the two papers published by Darsee in the *New England Journal of Medicine* (in one case with two coauthors and in the other with only one), "the two formal retraction notices, as well as Darsee's supporting letter, . . . seem to suggest that his coauthors at Emory had no responsibility at all for what happened, simply because they are honest and had no hand in the manipulation of the data. I cannot agree, and neither will most other editors" (*7*).

We believe that Sigma Xi should make clear that it shares Relman's view, through several actions taken by the Society and its members:

5. Sigma Xi could adopt principles that are similar to those suggested by Broad and Wade or by Relman, and that embody the responsibility in scientific publication the Society expects of its members.
6. *American Scientist* could make explicit what it already assumes implicitly: that it expects authors to submit papers in accordance with these principles.
7. Members (and, for that matter, nonmembers) of the Society could be encouraged to use the phrase "this paper is published in accordance with Sigma Xi's Principles" in the text or footnotes, and scientific journals could be asked to permit, encourage, and make well known these principles and such citation of them.

Publish or perish

One undeniable incentive for multiple authorship, and for scientific fraud, is the "publish or perish" attitude that tends to dominate scientific careers, at least in universities. It cannot wholly explain, let alone justify, either phenomenon, but it has other pernicious effects in science besides these. It encourages excessive specialization, so that younger scientists rapidly forget (if they ever learned) that major progress may come as much through cross-fertilization of one branch of science by another as it does through knowing more and more about less and less. A publish or perish mentality gives little incentive to teaching the research scientists of tomorrow, nor does it encourage those with a still broader vision of the scientist's task—a vision that might, for example, embrace support or leadership of a Sigma Xi chapter or club.

If such an approach to scientific research can be justified anywhere (and this seems unlikely), it can only be at a handful of institutes of advanced study where research is the dominant or only task. But, as the editors of *Nature* have argued recently, "the most serious pressure to publish comes from less distinguished institutions, determined to ape their betters and to enforce as rigorously as they can the most demanding

objective criteria that they know of" (*8*). We find the word "objective" ironical in this context, since it is precisely because the objective test of quantity is used in place of the subjective test of quality that publish or perish can become—and perhaps has already become—an evil that threatens the integrity and future of scientific research.

There is, of course, no particular virtue in infrequent publication as compared to a long list of research papers, though it might be argued that a vital element of good research is time for reflection, both before and after publication. Good scientists may publish a lot or a little. But there is a very definite evil in a university that allows or encourages tenure committees to set standards of, say, twenty published papers or abstracts in four years as a minimum requirement for consideration, or to discard as irrelevant any paper in a branch of science other than the tenure candidate's principal field of specialization. At a different level, for review panels of the National Science Foundation, National Institutes of Health, or similar bodies to assign major research funds on the basis of the quantity of material previously published by a laboratory is to accept a Faustian bargain that exchanges the long-term future of a national scientific research effort for a speedy decision and the next plane home.

If fraud is rare, and irresponsible authorship limited at present mainly to certain areas of science, publish or perish is ubiquitous across disciplinary boundaries, and has no regard to the character of an institution. It has been intensified by the drop in the birthrate, which has produced a highly competitive academic environment in which employment opportunities are created mainly by death, retirement, or failure to achieve tenure. There is little prospect of a self-correcting mechanism. Few institutions have the power or inclination to resist; the individual scientist, especially at the outset of a career, usually has neither.

The relevance of this issue for Sigma Xi is obvious. We are those who must exist in a publish or perish environment. We are the individuals who face tenure committees or grant-review panels that value quantity over quality. We are also the ones who sit on those committees or panels and who establish the criteria; the publish or perish mentality has been created by us. "Peer review" was established to be a safeguard for the quality of scientific research and for the freedom of the individual scientist. It is difficult to argue (even though the alternatives may be worse) that peer review is nowadays any more effective at maintaining these values than it is in reducing scientific fraud. Yet it could be effective in all of these.

If Sigma Xi is to combat the publish or perish mentality its actions might include the following:

8. Incorporation in the Society's Constitution of a clear statement that quality is the appropriate measure of the value of scientific research and that quantity is not.
9. Action by individual chapters and clubs to ensure that tenure committees and similar bodies adopt standards and procedures that recognize quality of research rather than quantity.
10. Use of Sigma Xi's influence among funding bodies such as NSF and NIH to ensure that similar standards and procedures are maintained and monitored.
11. Preparation by Sigma Xi of model standards and criteria that can be adopted by other institutions.
12. Actions by individual members, especially those in a position to influence the scientific careers of others, that work to eliminate a publish or perish environment and that reward quality research.

An unavoidable responsibility

At a meeting to make preliminary plans for Sigma Xi's centennial, it was suggested, not altogether frivolously, that the Society should survey the scientific environment that has emerged over the past 100 years, declare that Sigma Xi had succeeded in its objectives, and dissolve itself. It seems evident to us that the problems existing in this environment challenge an honor society for research scientists on a scale that Sigma Xi's founders could not have anticipated. It may be that, if Sigma Xi neglects these issues, they will be addressed effectively by others. But surely it cannot be denied that scientific fraud, irresponsible authorship, and an obsession with quantity over quality are vital issues for Sigma Xi and its members. The honor of scientific research is threatened, and our honor society cannot avoid its responsibility.

References

1. W. Broad and N. Wade. 1982. *Betrayers of the Truth*. Simon and Schuster.
2. The discussion on "Fraud and dishonesty in science" was reported in *Nature* 303:369 (2 June 1983) and in *New Scientist* 98 (9 June 1983). Tapes of the discussion can be purchased from Mobiltape Co., Glendale, CA.
3. A. S. Relman. 1983. Lessons from the Darsee affair. *New Engl. J. Med.* 308:1415–17.
4. Ref. *1*, p. 221.
5. Ref. *3*, p. 1417.
6. S. Hodge and D. Greenberg. 1981. Letter to the editors. *Science* 213:950.
7. Ref. *3*, p. 1417.
8. *Nature* 303:361. 2 June 1983.

Reported Laboratory Frauds in Biomedical Sciences

In recent years, we appear to have been hit by a wave of accounts in journals and newspapers of fraudulent reporting of scientific data. Many factors may have contributed to this, including an increase in the number of persons devoting substantially full time to research; an increase in the competition for available support dollars; an increase in the pressure to publish abundantly; possible changes in the nature of the training process that leads to the Ph.D. degree; and possibly, just possibly, a deterioration in the morality of our scientists. Assignment of weights to each of these and other factors is, at this time, not possible. However, review of reported cases indicates that certain generalities may be noted.

The frauds appear to occur most frequently in our most prestigious research and teaching institutions. A young man enters the laboratory of a very successful investigator, one who has an unusually large bibliography and who has, therefore, secured generous grant support and abundant laboratory space. The young man (there is a paucity of cases in which women are involved in fraud) inspects his new environment to ascertain what it takes to succeed there, and he soon concludes that, since his preceptor, who is obviously successful, has published an unusually large number of papers, this is the route to success. He therefore tries to follow this example and may publish ten or more papers in a year. Those of us experienced in the production of biomedical research know that to generate novel research results sufficient to fill that number of papers in the allotted time is not easy. Our young candidate, therefore, is forced by circumstances to consider the routes he may follow to achieve his ambition. On the one hand he may choose to plagiarize from the results of others. On the other hand he may choose to fabricate results and experiments. By these means, he hopes to diminish the time normally consumed in honest research, which entails the planning and design of experiments, the acquisition of the necessary technical skills, consideration of the relevant literature, and finally the preparation of his own research report. All of these steps consume time and energy.

One of the consequences of the selection of a very successful investigator as preceptor is that he will have attracted the attention of other pre- and postdoctoral fellows. He is likely to have a large number of trainees at some level of training who are accountable to him. Sometimes this number exceeds 20, and this, together with the fact that successful preceptors are likely to spend much time on the road giving lectures and participating in various committees and editori-

al board functions, reduces the time the preceptor may spend with each trainee. There are cases in which the trainee meets with his preceptor for 1/2 hour every 6 weeks. This is simply not enough if the preceptor is to take his teaching responsibility seriously. He should review with each trainee experiments done yesterday, and he should outline with the trainee the experiments planned for tomorrow. He should guide the trainee through the maze of the literature, and he should assist the trainee in the acquisition of the necessary experimental skills.

In addition, both by example and by precept, he has an obligation to make certain that each of his trainees is fully sensitized to the absolute requirement of total honesty in the reporting of scientific results. Many professions recognize unacceptable deviations from strict honesty. The banker cannot tolerate the embezzler. The military person cannot tolerate the deserter. And the intelligence services cannot tolerate the mole. Science cannot tolerate the man who takes lightly his moral obligation to report strictly what is true.

In light of the foregoing, it is suggested that research training groups under a single preceptor be kept small in size, permitting abundant contact between preceptor and trainee. The preceptor, if he is to fulfill his moral obligation, will undertake to spend significant periods of time at frequent intervals with each of his trainees. The candidate trainee will be well advised to select a preceptor who is not excessively encumbered with large numbers of trainees. He should recognize that a preceptor with an excessively lengthy bibliography may set quantity of publication above quality of research and may thus be a poor role model. In several of the recently publicized instances of laboratory fraud, the trainee has been the prime target of criticism, while the preceptor has been treated sympathetically and with commiseration. There may be instances in which this distribution of blame is not appropriate. Frequently, when one member of a group produces many more publications than do his contemporaries, this is taken as an indication of unusual ability. It may be appropriate to view such exceptions as a basis for suspicion and an indication for scrutiny.

DeWitt Stetten, Jr.
National Institutes of Health,
Bethesda, Maryland 20205

OCCASIONAL NOTES

SPEED

An Essay on Biomedical Communication

A COLLEAGUE was recently asked by the editor of a prestigious general scientific journal to review a paper on a trendy topic within 24 hours. The request was made on a Monday, and my colleague was instructed to report his opinion by telephone, so that the paper, if acceptable, could be in press by Friday. He reviewed the paper and told the editor that it was interesting but could be markedly improved by inclusion of additional information. The editor, however, was unwilling to delay publication of the paper and published it as it was within a week.

A second colleague presented a lecture on what he considered to be a derivative aspect of a fashionable topic at a scientific meeting. After the presentation, a journal editor promised instantaneous acceptance of a paper based on the lecture and submitted to his journal. My colleague acceded to the invitation, and the paper was published in the next issue of the journal.

As editor of a biomedical journal, I am sensitive to the pressure for rapid publication and clearly read the message that speed is perceived to be very important. In fact, to stay competitive, my journal is one of a growing list that has a special category of "rapid publications." Authors may submit short papers that they believe to be of special importance, to be considered for rapid publication. Reviewers of these papers are asked to communicate their opinions within 72 hours by telephone. The logistics of arranging such reviews are extraordinarily complex, and the resulting quality of refereeing and accuracy with which criticism is doc-

tumented can be marginal. Even for papers not submitted for rapid publication, the editorial office staff telephones prospective reviewers so that time is not lost in futile mailings when reviewers are unable or unwilling to referee papers. The staff also telephones reviewers again and again when they fail to return reviews within a designated time. Nevertheless, delays inevitably occur, and the staff endures the abuse of authors who are frantic that their papers remain in editorial limbo. The costs of the telephone calls, special postage, couriers, and other schemes to speed the processing of papers are considerable.

As newspaper and magazine readers know, publication can take place overnight. Most biomedical journals, however, have traditionally published papers within two to six months after their acceptance. The factor limiting the time from acceptance to publication is the money a periodical is willing to allocate to copy editing, typesetting, printing, binding, and distribution. These activities take time and can be accelerated only by hiring more copy editors, typesetters, and other workers than most journal budgets allow. Journals with large circulations and corresponding revenues can afford to invest in speedy publication, although until recently, even for such journals, an interval of less than three months between acceptance and publication was exceptional.

What is the hurry that justifies such trouble and expense? Since formal publication of data as a journal article is the basic currency of scientific communication and the validation mechanism for medical advances, the most defensible answer, at face value, is that the faster scientific findings are published, the sooner other scientists can use the information and the more quickly society stands to benefit from advances.

However, there is little evidence that the rate of publication today has a limiting effect on scientific and medical progress or on the public welfare. In the case of clinical research, the rapidity with which a given article can lead to changes in practice may depend on how clear-cut a proposed advance appears to be, how fixed the ideas of practitioners are, the number of other practice options available, and the accessibility of the proposed new diagnostic or therapeutic approach. The balance of these issues is rarely such that publication within a few weeks or less will truly benefit patients. In both clinical and basic research, advances are announced at scientific meetings and through a large network of informal communication.[1] Furthermore, by the time results are submitted to a journal for publication, the authors of the paper are usually already working on the next step of the project, and this work proceeds irrespective of whether the submitted paper is immediately accepted by a journal.[2] Finally, there is no evidence suggesting that the appearance of articles in journals is sufficient to cause additional scientific work to be carried out rapidly on the basis of the specific information contained within those papers. Few scientists or practitioners alter their direction rapidly. A contributing factor in this inertia may be the time it takes to find and read new papers. Despite increases in the speed of publication, the rate at which a human being can read has not changed. In fact, the very speed at which new papers appear may lower the probability of a given paper being read, although the growing interest of the popular media in medicine and science helps call attention to some papers in journals in which the media take an interest.

A more plausible and honest reason for the hurry is "priority." The surest way for a scholar to obtain the recognition of peers (and more and more, the public) is to establish and personalize a claim to discovery before anyone else. Squabbles among scientists over the priority of discovery are an integral part of science past and present.[3] If some piece of research becomes the seminal contribution of a major scientific advance — an improbable occurrence for most researchers — perhaps the author will receive wondrous rewards. In previous years the winnings were primarily fame among scholars and academic promotions. Nowadays, commercial profits and recognition by the public are added inducements.

The Watson and Crick story, popularized in James Watson's *The Double Helix*[4] and internalized by today's scientists in their formative years, may play a part in this somewhat irrational quest for success. Anyone reading the book before gaining experience in research is given the impression that the elucidation of DNA's structure was a sporting event in which speed brought incalculable glory to the winners, that a single shot can win the battle for kudos in science, and that the failure to fire it first means loss of both the battle and the war.

Certainly, the concern that someone else will publish an article first about something a researcher is working on is justified. More than 60 per cent of scientists have had the experience of being anticipated or "scooped" in this way at least once during their careers, and 17 per cent more than twice.[5] However, many if not most scientists recognized for major accomplishments have systematically completed a large body of work in a relatively leisurely manner over a period of years. Many venerated scientists either were slow to publish or, more commonly, failed to obtain recognition until long after publication of their work. Some examples from the past and present are Charles Darwin, Hans Krebs, and Barbara McClintock. From this fact, it follows that being scooped in any one project does not necessarily preclude eventual recognition for scientific accomplishment.

To determine the extent to which rapid publication has been a factor in the success of a few leading scientists, I polled 22 prestigious and currently active biomedical scientists concerning their views about rapid publication. The group included 11 Nobel laureates and 16 members of the National Academy of Science, and all had received other major awards for research accomplishments. All 22 scientists replied, and 15 in-

dicated that rapid publication had little influence on their rise to scientific prominence. Four stated that rapid publication had been of moderate importance, and three allowed that it had been very important. Although personal accounts concerning research behavior may lack objectivity,[6] all six of the scientists polled whose publication practices involved frequent use of journals that publish articles very quickly, suggesting that rapid publication was important to them, were those who stated that rapid publication played a large or moderate part in their success. None of the researchers who considered rapid publication unimportant published frequently in such journals. Thus, for a majority of this sample of prominent scientists, rapid publication was not perceived to be the route to such prominence.

Most young scientists may not be aware and older scientists may forget how frequently great discoveries emerge stepwise from the work of several researchers sequentially or simultaneously; the point at which a discovery is consolidated is not always obvious,[7] and the assignment of priority is often difficult. Good investigators tend to make multiple discoveries.[8] Novices may also not know that in science, as in other endeavors, "haste makes waste." Pseudodiscoveries and false breakthroughs often melt away when subjected to a little additional thought or experimentation. Prematurely published bad science can be a discredit to its perpetrators and a dangerous nuisance to the rest of the world. Furthermore, the young investigator has more to lose by being wrong than the established researcher.

Returning to *The Double Helix*, the fact that Watson and Crick found the "right" answer to DNA's structure, publicized as a short paper (published less than three weeks after its submission), remains far more indelibly in the memory than the fact that their competitor, Linus Pauling, had previously published a wrong conclusion in an unreviewed article. Moreover, with Pauling thinking he had already solved the problem, it is unclear who would have beat the heroes to the finish line, had the process been dragged out a bit. If we accept Crick's later assessment of this question, the answer is that there was no hurry.[9]

In exploring the matter further, we might also ask whether a perceived need for rapid publication is a consequence rather than a cause of outstanding achievement in science. The solution to a major problem in science often begins with the introduction of a new method, the discovery of a new phenomenon, or the proposal of a new conceptual approach, which then opens a pathway of derivative ideas and suggests experiments that lead to the problem's ultimate solution. It is these obvious subsequent steps in problem solving that ordinary scientists can complete speedily. As eloquently stated by Ziman, "Obsession with sheer immediacy and speed is a neurotic symptom, mainly to be observed in certain fashionable fields where too many people are chasing too few ideas."[10] The hidden steps, or the steps requiring extraordinary effort, demand the talents and usually prolonged efforts of extraordinary (or extraordinarily lucky) scientists.

A desire to have one's completed work published within a reasonable time is appropriate, and an occasional hastening of publication may be justified. However, what is now a cause for alarm is the growing trend of the scientific media to feed on the community neurosis about speed and escalate the emphasis on rapid publication. Now, the journal that can promise rapid publication has a competitive edge. Can editors be trusted to decide what findings truly deserve such extraordinary acceleration?

Ostensibly organized to serve scientists and clinicians, journals also have a commercial prerogative to survive. Journals garnering large readerships that in turn attract advertisers can generate substantial revenues and survive in grand style. To the extent that publication of the most fashionable or newsworthy papers promotes circulation, it would not be surprising to find editors using the prospect of rapid publication as bait to attract such fare. The "arms race" for trendy research that has begun to take place among some of the most prestigious journals, as epitomized by the incidents described in the beginning of this essay, has serious implications for the research community. First of all, it "normalizes" the neurosis by increasing the pressure for more journals to accelerate their rate of paper processing. Secondly, it means that the journal editors begin to redefine what is normal behavior among scientists. Thirdly, it imposes more onerous burdens on the community of scientists asked to perform peer review. In addition, this practice may encourage the submission of incomplete work, which contributes to fragmentation and pathologic hypertrophy of the biomedical literature. Finally, it undermines the value of prepublication forums for determining priority and providing a preliminary review of scientific work.

In the past, a well-documented public presentation of a research result at a scientific meeting, accompanied by a published abstract, constituted an adequate priority claim, if followed eventually by a full journal article. The increasing willingness of journals to accommodate nearly instantaneous publication of results, which constitutes an immediate formalization of priority, threatens not only to negate the impact of meeting abstracts but indeed to promote their extinction. If there is no value in publishing meeting abstracts, then scientists may simply view them as "intelligence leaks" to competitors and refrain from submitting them. Thus, the pressure for rapid publication could hasten the obsolescence of scientific meetings at which scientific work has traditionally received initial criticism and refinement. If present trends continue, presentations at such meetings will increasingly serve the limited purpose of marketing already published findings.

Supposedly, the peer-review system is an instrument used by scientists to ensure quality control and protection of their claim of discovery. However, the extent to which researchers themselves control this system is unclear. It is believed that criticism by referees and the balance of competing ideas and topics offset the normal and appropriate passion of researchers to promote their own work vigorously.[11] However, journals, especially those with large readerships and liberal budgets, have the power not only to select competing papers for publication but also to editorialize ("hype") their content. Whether, as Feyerabend has suggested,[12] this practice influences the opinion of the readership and creates an image of outstanding science where none in reality exists is unclear, but it certainly helps authors promote their research. If journals use speedy publication as a weapon in the competition for fashionable research, they play on the vanity of scientists and distort the system of values. Not only can they decide what type of work will receive favored treatment, but they overburden expert reviewers by demanding more rapid reviews. Furthermore, they expose the reviewers to a conflict of interest. If a competitor's work in the same field is worthy of publication within a week, then by implication, the reviewer's work merits the same treatment. To refuse to review, to review negatively, or to downgrade the need for rapid publication is to risk lack of reciprocal treatment.

If rapid publication has been unnecessary for the ultimate success of many researchers, wide appreciation of the fact could lead to some self-examination and behavior modification. Investigators themselves set the standards for rewarding their peers. If the majority of highly rewarded scientists had little need to resort to speed in publication, everyone else might be inclined to relax a bit. The biomedical community would be spared costs and might actually gain some time to think and work.

Massachusetts General Hospital
Boston, MA 02114

THOMAS P. STOSSEL, M.D.

REFERENCES

1. De Solla Price DJ. The scientific foundations of science policy. Nature 1965; 206:233-8.
2. Garvey WD. Communication: the essence of science. Oxford: Pergamon Press, 1973:125.
3. Merton RK. Priorities in scientific discovery. Am Sociol Rev 1957; 22:635-59.
4. Watson JD. The double helix. New York: Atheneum Press, 1968.
5. Hagstrom WO. Competition in science. Am Sociol Rev 1974; 39:1-18.
6. Gilbert GN, Mulkay M. Opening Pandora's box: a sociological analysis of scientists' discourse. Cambridge: Cambridge University Press, 1984.
7. Kuhn T. The structure of scientific revolutions. 2nd ed. Chicago: University of Chicago Press, 1970.
8. Merton RK. Singletons and multiples in scientific discovery: a chapter in the sociology of science. Proc Am Philos Soc 1961; 105:470-86.
9. Crick F. The double helix: a personal view. Nature 1974; 248:766-9.
10. Ziman JD. Information, communication, knowledge. Nature 1969; 224:318-24.
11. Polanyi M. Personal knowledge: towards a post-critical philosophy. Chicago: University of Chicago Press, 1962.
12. Feyerabend P. Consolations for the specialist. In: Lakatos I, Musgrave A, eds. Criticism and the growth of knowledge. Cambridge: Cambridge University Press, 1970:197-230.

Publish or Perish: A Proposal

MARCIA ANGELL, M.D.; Boston, Massachusetts

Because promotion and funding of physicians in academic medicine are closely linked to the number of their publications, investigators feel impelled to publish as frequently as possible. This pressure leads to a number of unfortunate practices in medical publishing, including undertaking trivial studies because they yield rapid results, needlessly reporting the same study in installments, reporting a study more than once, and listing as authors people only marginally involved in the study. It may also be a motivation for fraud. An effective way to reduce these offenses and affirm the supremacy of substance over volume in scientific research would be to place a ceiling on the number of publications that can be considered in evaluating a candidate for promotion or funding. Each publication would then receive commensurately more attention, both from the researcher and from those judging the work.

PHYSICIANS IN academic medicine are evaluated for promotion and funding in large part according to the length of their bibliographies—a fact that is well understood in the academic community, despite occasional protestations to the contrary. A minimum number of publications is usually necessary, if not sufficient, for promotion on both the academic and clinical ladders. Thus, good teachers and clinicians who have little interest or ability in research must at least go through the motions of research and publication in order to advance. Renewal of funding also depends on publications as a reflection of productivity. Not surprisingly, researchers feel impelled to publish as frequently as possible. Some resultant practices, such as publishing in "least publishable units" (1), publishing the same study in different forms, or adding supernumerary authors to articles, are so common as to be almost standard. A researcher may also choose work that yields rapid and, if possible, positive results, rather than long-term uncertain studies that may be scientifically more important.

There are many drawbacks to rewarding large numbers of publications. The practices of fragmented and repetitive publication are wasteful and inefficient. They also produce a huge and unwieldy literature that is difficult to evaluate because of its size and the clutter of repetitious and often trivial reports. The bias toward studies that can be completed and published relatively quickly distorts the direction of science. Whether the pressure to publish leads to major fraud, such as plagiarism and fabrication of data, is difficult to say. Persons who are predisposed may engage in fraud for this reason, but motivations in these cases are far from clear. Lesser offenses in medical research, however, ranging from repetitive publication to the selective reporting of data that support the hypothesis, clearly stem from the pressure to publish with its almost irresistible incentive to cut corners. I believe that these practices, too, entail some deception (2), even though they are not what we usually mean by fraud. For example, a duplicate publication is usually not identified as such by the authors; instead, it is often inflated as a unique contribution. Thus, these practices constitute an erosion of the integrity of the scientific enterprise and debase its coinage.

On the other hand, there are advantages to the present system. The length of a bibliography is, after all, a relatively simple and unambiguous criterion for gauging output in a large and complex enterprise. More important, our system virtually ensures that scientific work is reported in published form as rapidly as possible. Nearly any study, no matter how preliminary or inconsequential, can probably be published somewhere. Thus the exchange of information in our system tends to be rapid and complete. The pot is always boiling, and this sort of energy is good for science. Talented researchers who, if left to their own predilections, may be slow in getting things done have an incentive to produce and make their findings known.

How can the system be reformed to eliminate its flaws while maintaining its strengths? Huth's (3) suggestions that scientific organizations, institutions, and journals try to develop standards to discourage fragmented and repetitive publication and to define criteria for authorship constitute valuable first steps. But as long as the pressure to publish remains, these steps would address the effects of the problem and not the cause. The source of the problem must also be addressed, both at the institutions where research takes place and at the funding agencies, in particular, the National Institutes of Health. Institutions that now pay lip service to the importance of excellence in teaching as a criterion for promotion should begin to make serious attempts to evaluate it and, when appropriate, give it the weight they now attach to research. Perhaps excellent clinical teachers should not be required to maintain a veneer of research capability, but should be evaluated entirely on the basis of their teaching and clinical work, as suggested by Petersdorf (4). The short cycles of the funding agencies, with the attendant need to produce immediate results, adds to the pressure to publish and, in addition, discourages long-term studies. Increasing the length of the cycle, at least in some cases, or awarding long-term grants for promising investigators may help to eliminate this problem.

These reforms would be helpful. I believe, however, that the single most important reform would be to set a ceiling on the number of publications that may be consid-

▶ From *The New England Journal of Medicine*, Boston, Massachusetts.

©1986 American College of Physicians

ered for promotion or funding. I suggest that any institution or agency evaluating a researcher for promotion or for funding consider only, at most, the three articles the candidate considers to be his or her best in any given year, with a maximum of perhaps ten in any 5-year period. Other publications should not even be listed. Each publication would then receive commensurately more attention, both from the researcher and those evaluating the work. At present, it is nearly impossible for those evaluating a researcher who has a long bibliography to become acquainted with all of the work. After all, there may be a hundred or more papers. If the number were limited, the emphasis would shift from the quantity of the research to its quality, and it would become both necessary and possible to evaluate each study.

How would such a system affect questionable practices and fraud in medical research? Almost certainly the lesser offenses described by Huth (3) and Bailar (5) would be curtailed, and our large and largely trivial literature would become leaner and more substantial. There would no longer be an incentive to test the lower limits of the "least publishable unit," to publish the same study in several different disguises in several different journals, and to confer authorship on people who were not authors by any reasonable criteria. The frequency of these and other relatively minor forms of deception would almost certainly be greatly reduced. At the same time, more ambitious studies would be undertaken and not published until they were completed. I believe that most researchers would welcome the opportunity to turn their attention more fully to the substance of their work without having to calculate in advance the number and timing of the publications that might emanate from it.

Whether this reform would reduce the frequency of the more egregious forms of fraud—plagiarism and fabrication of whole bodies of data—is problematic. It is possible that these forms of deception would not be much affected. For persons disposed to fraud, the decrease in the pressure to publish as an incentive might be offset by the increase in the pressure to publish something of substance. After all, it is more difficult to complete a good study than a trivial one. But it would probably be easier to uncover fraud in a system that focused more deliberately on the quality rather than the quantity of work.

In summary, limiting the number of papers considered for promotion or funding would have three probable effects, as I have suggested elsewhere (2). First, the quality of medical research would tend to improve, insofar as each study would receive commensurately more attention. Second, promotions and funding would more accurately reflect the quality of a researcher's work, because a smaller number of publications would be easier to evaluate. Third, some of the fluff in our huge scientific literature would be eliminated. I believe this reform would also decrease the prevalence of many questionable practices in the system, although it might not affect the frequency of major fraud. In the end, the most important result would be to affirm the supremacy of substance over volume in scientific research.

▶ Requests for reprints should be addressed to Marcia Angell, M.D.; *The New England Journal of Medicine*, 10 Shattuck Street, Boston, MA 02115.

References

1. Broad WJ. The publishing game: getting more for less. *Science*. 1981;211:1137-9.
2. Angell M. Editors and fraud. *CBE Views*. 1983;6:3-8.
3. Huth EJ. Irresponsible authorship and wasteful publication. *Ann Intern Med*. 1986;104:257-9.
4. Petersdorf RG. Is the establishment defensible? *N Engl J Med*. 1983;309:1053-7.
5. Bailar JC. Science, statistics, and deception. *Ann Intern Med*. 1986;104:259-60.

Science, Statistics, and Deception

JOHN C. BAILAR III, M.D., Ph.D.; Washington, D.C.

Some common scientific practices cannot quite be called lying, though they are potentially, and sometimes deliberately, deceptive. Some examples are the failure to explain to readers all the weaknesses in data, statistical testing of *post hoc* hypotheses, fragmentary or selective reporting of findings, and reporting as "negative" a study that had insufficient chance of detecting an effect. The first step toward controlling potential problems is a redefinition of ethical standards to bar readily avoidable as well as deliberate deception. Other remedies include substantially greater restraint in the use of questionable practices, full disclosure and justification each time these practices are used, and greater skepticism by readers. Pressures to publish tend to promote deception. A broadened concept of ethical standards in science should be reflected in training programs and in the structure of scientific rewards, including a sharply reduced emphasis on publication as an end in itself.

IN SCIENCE, lying is condemned, even by some of its few practitioners. Deliberate or careless deception short of lying, however, seems to be universally accepted and sometimes even promoted as a part of the culture of science. I do not suggest that scientists as a group are careless, venal, or otherwise depraved: they may even be above the human average in developing and adhering to detailed, albeit tacit, standards of professional conduct. Those who are clearly violators are drummed out of our ranks, loudly and publicly. But what about less clearcut deception?

My thesis is that our professional norms are incomplete and that several kinds of widely accepted practices (Table 1) should also be widely recognized as potentially deceptive and harmful. Some of these practices also have much value, but at times they are inappropriate and improper and, to the extent that they are deceptive, unethical.

The scientific method is fundamentally concerned with the processes of inference, generally from data that are necessarily inaccurate to some degree, incomplete, drawn from small samples, or not quite appropriate for a specific task. Inference—that is, drawing conclusions or making deductions from imperfect data—provides most of the excitement and intellectual ferment of science. Scientific rewards are probably more closely related to valid inferences established than to such related activities as imaginative hypotheses formulated, elegant experiments designed and conducted, or new methods developed. The rewards for publishing a first-class inference can include income, position and power, professional status, and the respect of colleagues. Such rewards may sometimes count for more than self-respect and the joy of discovery. We must therefore be attentive to scientific norms and activities that may distort the processes of inference.

An example of a deceptive practice is the statistical testing (such as the calculating of p values) of *post hoc* hypotheses. It is widely recognized that t-tests, chi-square tests, and other statistical tests provide a basis for probability statements only when the hypothesis is fully developed before the data are examined in any way. If even the briefest glance at a study's results moves the investigator to consider a hypothesis not formulated before the study was started, that glance destroys the probability value of the evidence at hand. Certainly, careful and unstructured review of data for unexpected clues is a critical part of science. Such review can be an immensely fruitful source of ideas for new, before-the-fact hypotheses that can be tested in the correct way with other new or existing data, and sometimes findings may be so striking that independent confirmation by a proper statistical test is superfluous. Statistical "tests" are also used sometimes in nonprobability ways as rough measures of the size of an effect, rather than to test hypotheses. (An example is the column of p values that sometimes accompanies a table comparing the pretreatment characteristics of patient groups in a randomized clinical trial.) When either the test itself or the reporting of the test is motivated by the data, a probability statement such as "$p < 0.05$" is deceptive and hence damaging to inference.

Other potential problems are the selective reporting of findings and the reporting of a single study in multiple fragments. These practices can obscure critical aspects of an investigation, so that readers will misjudge the evidential value of the data presented. Such reporting may be deceptive, whether deliberately or accidentally. On the other hand, these practices sometimes have positive value that should be preserved. For example, they can facilitate the tasks of both the investigator and the user when a demand for a monolithic analysis might seriously delay or frustrate the progress of both.

"Negative" conclusions of low statistical power—that is, reporting that no effect was found when there was little chance of detecting the effect—can also distort inference, especially when investigators do not report on

▶ From the Office of Disease Prevention and Health Promotion, Department of Health and Human Services, Washington, D.C.

Annals of Internal Medicine. 1986;104:259-260.

Table 1. Some Practices That Distort Scientific Inferences

Failure to deal honestly with readers about non-random error (bias)
Post hoc hypotheses
Multiple comparisons and data dredging
Inappropriate statistical tests and other statistical procedures
Fragmentation of reports
Low statistical power
Suppressing, trimming, or "adjusting" data; or undisclosed repetition of "unsatisfactory" experiments
Selective reporting of findings

statistical power. The concept of power is formally defined in terms of the random variability of results that is inherent in a specific combination of data structure, sample size, statistical models, and analytic method; but I believe that the concept should be substantially broadened to include the likelihood that a particular effect would be detected and reported if it were present to some specified degree. Such an analysis rarely accompanies "negative" findings, and readers may be left with an unjustified sense that an effect not demonstrated is an effect not present. Again, however, there are counterarguments: a report with low power may be better than no report (and no power), or meta-analysis (1) of several low-power reports may come to stronger conclusions than any one of them alone. Reporting negative studies of low power can create ethical problems, but those problems may be largely mitigated if the low power is accurately and clearly reported as well. Too many scientists resist the objective reporting of this kind of weakness in their work, and pressure for "strong" results may be greatest during the formative years of graduate training and career entry. Thus we may be training new scientists in unethical methods.

Despite the occasionally useful roles of these and other practices listed in Table 1, each can seriously distort the processes of inference and should therefore be an object of concern. Where the practices have legitimate applications, they should, of course, be used; but even then they should be fully and explicitly disclosed by the investigator, justified in some detail, and accepted with caution by readers. A combination of restraint in their overall use, limiting their use to clearly appropriate situations, providing full disclosure and justification, and maintaining the readers' skepticism will help to diminish the frequency and severity of ethical problems. Full disclosure here means more than a few words buried in the fine print of a Methods section; it means not just that the author send a message, but that the author also work to ensure that the message is received and correctly interpreted by readers. There are parallels here to the evolving requirements for informed consent by experimental subjects.

Pressures to publish can be great and may account for many of the abuses suggested by Table 1. I fear that even the constrained use of these potentially damaging practices will leave attractive loopholes for an army of ambitious practitioners of science, each feeling great pressure to publish, who will rush in to explain why his or her situation is different, why full disclosure is inappropriate, and so forth. I am convinced that science, scientists, and society as a whole would benefit from substantially broader concepts, ultimately based on the need to protect the processes of inference, about ethical standards and violations in science.

▶ Requests for reprints should be addressed to John C. Bailar III, M.D., Ph.D.; 468 N Street, S.W.; Washington, DC 20024.

Reference

1. Louis TA, Fineberg HV, Mosteller F. Findings for public health from meta-analysis. *Annu Rev Public Health.* 1985;6:1-20.

Pressure to Publish and Fraud in Science

PATRICIA K. WOOLF, Ph.D.; Princeton, New Jersey

Pressure to publish is an overworked and ill-used phrase that encompasses at least two meanings: the pressure an individual scientist may feel, and the pressure exerted by the institutions of science. It is important not to confuse the two. A modest level of institutional pressure is legitimate and healthy for science. Persons who respond to these pressures by publishing extravagantly may be doing a disservice to impressionable young colleagues and to the notions of responsible authorship.

Q. Is there pressure to publish in today's science?
A. Yes, there is.

Q. Is pressure to publish bad for science?
A. Bad or good, we're stuck with it. Since its birth, modern science has had an ethos that called for communications; scientists have felt some pressure to make their results known. Publication is an established way of doing so.

Q. Should there be pressure to publish?
A. Why not? If the public has paid for research, they are entitled to evidence that research has been done.

Q. Are publications the product of research? Doesn't the public pay for new knowledge and ideas?
A. Yes, that's the ideal, but not all experiments work. The public is really paying for a good-faith effort by highly skilled people who collectively have a pretty good chance of coming up with some information that is an advance.

Q. How do we know which scientific results represent advances in scientific knowledge?
A. That's where publication comes in; information known only to the discoverer cannot be evaluated and become knowledge. Some say it isn't even science until it is published. Research reports are evaluated by editors and referees before they are published, and then publication allows other scientists to further evaluate and use the results.

Q. Why is there so much concern about pressure to publish?
A. Many persons in science are worried that their important but complicated system of communication is in jeopardy. Because of undue pressure it may not be working the way it's supposed to.

Q. Why is it complicated?
A. Publication is no longer just a way to communicate information. It has come to be a way of evaluating scientists; in many cases it is the primary factor in professional advancement.

Q. Is that so bad? If articles are refereed, doesn't that ensure a breadth of professional judgment?
A. Some say that the quantity of published papers has become more important for evaluation than the quality of ideas they contain.

Q. Is that true?
A. It's hard to say. The evidence is anecdotal. Promotion and granting committees meet in secret; according to a dean in a major medical school, "Confidentiality fans the flames of paranoia." There are no systematic studies of these deliberations, but some reputable people have criticized the system. We need more systematic information about the relation of publications to professional advancement.

Q. Why are they so critical if the facts are not known?
A. Their experience has led them to be concerned about science. And they are worried that careerism is damaging a very important national resource—our research establishment.

Q. Is there that much pressure to publish?
A. It depends on what sort of pressure you mean.

▶ From the Department of Sociology, Princeton University, Princeton, New Jersey

©1986 American College of Physicians

Q. Is there more than one kind of pressure to publish?
A. There are at least two sorts: the pressure the scientist feels from the tradition of science, and the pressures from institutions.

Q. Are these pressures the same?
A. How could they be? Institutions and their departments may have different standards for promotion. Requirements are often implicit rather than explicit, and frequently they are misunderstood. Some research fields publish more heavily than others.

Q. Why have so many of the recently disclosed frauds been found in medical schools?
A. No one knows for sure. Their research environments are highly competitive. The heroic nature of their quest directly affects issues of life and death, so there are greater economic and social rewards in medical science. Another possibility is that some fraud occurs in all science, but it is more likely to be detected in a rapidly growing research area like biomedicine.

Q. Is there more pressure on medical scientists than on others?
A. Some say there is. Medical researchers often have clinical responsibilities in addition to teaching, research, writing, and applying for research funds.

Q. Have any studies measured psychological pressures on scientists?
A. I haven't found any in a search of the literature. One study currently underway at a major medical school is trying to find out if persons who are up for promotion really understand the promotion and tenure procedures of the institution.

Q. How can we measure the amount of pressure being put on scientists by institutions?
A. One way is to look at the amount of publishing done by practicing research scientists in our best institutions.

Q. Aren't you just counting papers again?
A. It's not the best way of determining research outcomes, but if becoming a full professor at a good university takes an exorbitant amount of publishing, we probably have a problem on our hands.

Q. How many papers does the average scientist publish?
A. A study by King and colleagues (1) showed that in 1977 the average number of articles published per physical scientist was 0.20/year, up from 0.17/year in 1965. The average number per life scientist in 1977 was 0.40/year, up from 0.30 in 1965. In this period only the social sciences showed a big decline; the average number of papers per social scientist was 0.22/year, down from 0.60 in 1965 (1). But these data include industrial or government scientists who may not be actively publishing. A better definition of scientist is probably a "publishing scientist." Price and Gursey (2) have determined that among publishing scientists, the average productivity is 1 paper per scientist per year.

Q. What are the rates of publication of scientific faculty in our major universities?
A. They differ among disciplines. In a study of the faculty of nine distinguished universities, rates ranged from 1.8 publications per scientist per year in physics to a high of 2.7 publications per scientist per year in biochemistry. Astronomy, chemistry, and microbiology were the other specialties studied. Publication records of over 1000 faculty members were examined (3).

Q. What about medical schools?
A. There are many medical schools whose faculty publish very little. In the 1973-1975 period the average rate of publication for privately controlled medical schools was 250 papers per year and for publicly controlled medical schools, 150 papers per year (4).

Q. What do federal granting agencies expect from their investments in research? Is there unreasonable pressure to produce simply to satisfy these requirements? On the average, how many papers per grant per year are produced?
A. The Office of Program Planning and Evaluation of the National Institutes of Health (NIH) has examined the literature output of NIH-supported research. A recent study of faculty in departments of medicine shows that one third of M.D. faculty are not significantly engaged in research at all (5). The NIH makes between 5000 and 6000 competitive grants per year. The average length of the grants is 3.2 years and that amounts to roughly 16 000 to 20 000 grants in force. In the same period and with the predicted lag of 3 years between a grant award and its resultant publication, 16 000 to 20 000 papers per year are produced. These figures are subject to a possible 20% to 30% increase, because at different periods some journals were not included in the survey. But that amounts to 1 or 1.3 papers per grant year (6; Gee H. Personal communication).

Q. If you double that to 2 or even 3 papers a year per grant, would that be too much pressure on a scientist?
A. It doesn't seem like it.

Q. If these institutional expectations are relatively modest, why is fraud associated with pressure to publish?
A. There are two principal reasons. First, several of the most spectacular frauds have taken place in laboratories where the number of papers published every year significantly exceeded the norm. Second, many of the persons involved have mentioned pressure as a factor. The newspapers usually report that view, and pressure has come to be a stock explanation.

Q. Could we look at those two factors separately?
A. That's sensible. "Pressure" has come to be a convenient explanation, for two reasons. The reports of these events have shown how shocked scientists are to discover that one of their colleagues was not playing by the rules. They may explain that fault using a psychological excuse that focuses on the flawed personality

of an individual person and not on science in general; it is a less severe condemnatory judgment than, for instance, moral turpitude.

Q. Is that what's called medicalizing the deviance?

A. Yes, it seems to isolate the problem and let the other scientists off the hook. But then questions began to be asked about where the pressure comes from.

Q. Does that bring us round to the publishing practices in laboratories where fraud was discovered?

A. Let's look at just a few. The first of the major recent scandals, that of William Summerlin at the Memorial Sloan-Kettering Cancer Center, was in the research group supervised by Robert A. Good. In the 5 years before 1975, Good had published 342 papers, an average of 68/year (16 as first author, 325 as coauthor with at least 136 coauthors). Several other frauds were associated with medical research groups that have been extremely prolific. In the period 1975 to 1980, which preceded the discovery of fraud in their laboratories, Eugene Braunwald had published 171 papers, an average of 29/year. Philip Felig had published 191 papers, an average of 31/year. Ephraim Racker, who is not on a medical faculty, had published an average of 16/year (7).

Q. How do these pockets of prolific publication compare with similar departments in similar institutions?

A. These scientists have made important contributions, significant far beyond the numbers of their publications. Therefore, comparisons with average productivity have limited value. But, yes, their average levels of publishing are significantly higher.

Q. Should we be worried about such prolific publishing?

A. There is reason for concern on several levels. Simple time-and-motion studies lead to the conclusion that when an author's productivity is extremely high, his participation in each paper has to be low. Or at least there is no clear way of telling which articles the author contributed to significantly. These prolific scientists are clearly operating in ways very different from those in the mainstream of science. One can argue that excellent science has always been out of the mainstream, but when there are significant deleterious effects on younger coworkers in their laboratories and eventual effects on scientific research in the United States as a whole, we have a responsibility to question the wisdom of those research and publishing practices.

ACKNOWLEDGMENTS: Grant support: from the Rockefeller Foundation.

▶ Requests for reprints should be addressed to Patricia K. Woolf, Ph.D., Department of Sociology, 2-N-2 Green Hall, Princeton University, Princeton, NJ 08544.

References

1. King DW, McDonald DD, Roderer NK. *Scientific Journals in the United States.* Stroudsburg, Pennsylvania: Hutchinson Ross Publishing Company; 1981:62.
2. Price D, Gursey S. Studies in scientometrics: Part I. Transience and continuance in scientific authorship. In: International Forum on Information and Documentation. Moscow: International Federation for Documentation; 1975:17-24.
3. Frame JD. Quantitative indicators for evaluation of basic research programs/projects. *IEEE Trans Eng Manage.* 1983;30:106-12.
4. McAllister PR, Narin F. Characterization of the research papers of U.S. medical schools. *J Am Soc Inform Sci. 1983;*34:123-31.
5. Beaty HN, Babbott D, Higgins EJ, Jolly P, Levey GS. Research activities of faculty in academic departments of medicine. *Ann Intern Med.* 1986;104:90-7.
6. Analysis of research publications supported by NIH and NIGMS. In: *NIH Program Evaluation Report.* Bethesda, Maryland: National Institutes of Health; 1980.
7. *Science Citation Index Source Index.* Philadelphia: Institute for Scientific Information; 1970-80.

PERSPECTIVE

Fraud, Irresponsible Authorship, and Their Causes

Abuses in scientific publication such as fraud and irresponsible authorship appear to be occurring more frequently. Whether this impression is due to abuses at a constant rate in a growing scientific establishment or to a decline in ethical standards in science is not known. But these abuses greatly concern editors and others with responsibilities for the integrity and reliability of the scientific literature. A panel arranged by Dr. Marcia Angell, Senior Deputy Editor, The New England Journal of Medicine, *for the meeting of the American Association for the Advancement of Science last May discussed these abuses. The five papers that follow are from the participants in that panel.*—The Editor.

The Pathogenesis of Fraud in Medical Science

ROBERT G. PETERSDORF, M.D.; La Jolla, California

Using a recent case of research fraud as a basis, this article examines the pathogenesis of fraud in medical science and suggests some remedies. Among causative factors are the "pre-med syndrome," the extraordinary size of science (which makes supervision of young investigators difficult), and competition, both professional and economic. Remedies include more careful selection of personnel, reduction of excessively large research groups, and closer examination of work at all levels—the laboratory, the academic department, and the institution. Each institution should have in place mechanisms to investigate research fraud when it is uncovered.

RECENTLY I had a disquieting experience. A young faculty member about to be appointed in a clinical department was discovered to have submitted several papers containing fraudulent data to several reputable journals. Up to this shocking discovery he had not been suspected of malfeasance. In fact, he was considered somewhat of a "child prodigy." He had done superbly in medical school and was excellently trained; during his fellowship he wrote over 30 papers. When he was proposed for his initial faculty appointment, he had authored in excess of 100 papers.

In assembling the proposal for this man's academic appointment, a senior member of the department reviewed some of these publications. The data of several control groups were the same in several papers: normally, different experiments should have different controls. The senior faculty member consulted with the coauthors. They had never seen the papers. Technicians in the laboratory could not recall the experiments. The data books could not be found. No purchase orders for the animals used in the experiments could be found. The grants to which the work had been credited had not contained enough money to purchase the animals used in the work that was published. When the young investigator was confronted with these doubts, he denied all the allegations but was unable to marshal proof that the allegations were false. His letter of resignation from the faculty followed shortly thereafter.

One might think that the episode ends here. Although a full investigation has corroborated the fraud in some papers, others need to be reviewed. The investigation will need to be continued: to protect the coauthors; to salvage the reputations of the research fellows who worked with the culpable investigator (but who can give them back the years of fellowship?); and to protect the granting agencies.

The "Pre-med Syndrome"

How have we gotten into this sorry mess? Believe it or not, it probably began with the "pre-med syndrome," painfully familiar to many in medical education. It begins with fierce competition in college, excessive emphasis on grades, and the rise of students who become 22-hour-a-day study machines. The syndrome was analyzed in detail in the May 1985 issue of *The Washington Monthly* (1). Some quotes from this article are instructive.

> A recent study of 400 medical students by doctors at two Chicago medical schools revealed that 88% of the subjects had cheated at least once while they were pre-meds. The researchers found widespread cynicism among their subjects as demonstrated by frequent agreement with statements such as "people have to cheat in this dog-eat-dog world."

The majority of students continued to cheat in medical school.

> The most disturbing finding was that there was a continuum from cheating in college to cheating in medical school in didactic areas to cheating in clerkships and patient care. Cheating in medical school may be a predictor to cheating in practice.

The culture in which we train our young promotes cheating.

▶ From the University of California, San Diego, School of Medicine, La Jolla, California.

©1986 American College of Physicians

63-471 - 87 - 19

From crib notes in organic chemistry to Medicare fraud and fudged results in medical research, the connection may be there.

The Large Size of Science

A second factor in the pathogenesis of scientific fraud is the very size of science itself. Despite some recent erosion of research budgets, they are much greater in absolute terms than ever before. The department of medicine that I chaired for 15 years, from 1964 to 1979, serves as an example (2). During this 15-year period the research budget quintupled while the teaching load only tripled. This increase in the research enterprise was accompanied by an increase in the size of graduate training programs. The number of housestaff tripled, the number of graduate students and postdoctoral fellows increased markedly, and the number of clinical fellows increased most of all.

In fellowship training groups in internal medicine, as many as 18 fellows in a group is not unusual. Despite claims that more than 50% of these clinical fellows go into academic medicine—claims that I have never believed even in the days of training grants—I have real doubts as to whether we need so many persons in advanced training, either for the academy or for practice. In fact, there is now good evidence that we do not need them for practice, where there is a surfeit of almost all types of specialists and subspecialists (3). The penchant for procreation among directors of clinical programs and among trainers of basic scientists knows no bounds; they feel that they have a right to generate the intellectual progeny of the future, whether that progeny becomes an overpopulation or not.

What are the consequences of big science? First, there is the inability to say grace over the large research groups that now exist. Thus these groups have an intermediate layer of junior faculty, many of whom are relatively inexperienced in training people and supervising research. It also results in group projects. There are not enough good ideas to go around to keep the excessive number of trainees busy. As a consequence, group projects abound with the inevitable result, the multiauthored paper. Characteristic of the multiauthored paper is that it is often not completely reviewed, and usually not in detail by every author. A few decades ago the department chair read every paper emanating from his department. This is clearly no longer possible, but now even division heads do not read the papers, as some in whose divisions fraud has occurred now ruefully admit.

Competition

A third factor in the pathogenesis of fraud is competition itself—not necessarily the competition to be first in press but, more often, the competition for grants. The progressive funding of grants with higher priorities does not connote that the science is necessarily better. It only means that competition for the research dollar is greater. Getting new grants or successful renewals requires productivity, and productivity, in turn, requires training of new people who can then publish more papers.

There is also competition for position, and our system of academic promotions is the prime example. Promotions committees count and weigh papers but do not read them. The reported research is usually narrow in scope, and the promotions committee quite often are unable to understand the nature of the work. It is much easier, therefore, for them to judge quantity not quality. Unfortunately, it is also far easier to judge quantity or quality of research than excellence in teaching and excellence in clinical care.

Academic personnel policies such as tenure exacerbate competition. Tenure appointments are guarded with jealousy and are made with a sense of false elitism, and tenure is often based on faulty criteria. Competition among young faculty for the scarce tenured slot returns the faculty to the days of the "pre-med syndrome." Tenure as a personnel policy in medical schools assures only a small fraction of the faculty member's income (4). The good do not need it and the mediocre use it as a shield. Tenure tends to perpetuate mediocrity.

Some academicians contend that the intense competition, both professional and economic, in science makes an occasional instance of research fraud inevitable. I do not share this fatalistic attitude. I believe firmly that the committed scientist cannot alter the truth. But even if some research fraud is inevitable, it is probably preventable, if not by the sponsoring institution, then by the principal investigator himself.

In this era of big science, principal investigators must be more careful in screening their associates. All too often younger investigators are invited to join a laboratory because the funds are available to support them or because they provide an extra pair of hands to do the work. Principal investigators must learn to contain the size of their laboratories. I have from time to time seen the size of a laboratory equated with the quality of work emanating from it. In science, as in all else, big is sometimes beautiful, but often it is not.

Original research data need to be checked at several levels; graduate students, postdoctoral fellows, and junior and senior faculty members must check and cross-check the raw data and not just the computer printouts, graphs, or figures that are often an intermediate product of research. Although data books may clutter the laboratory, they must be retained so that in case of doubt they will be available for review. Multiauthored papers require input and checking by all persons named as authors and not just the senior author. Heads of divisions, departments, institutes, and schools should be sufficiently familiar with the laboratories under their administrative purview so that they will be able to launch an investigation of alleged fraud when such an investigation is needed.

Conclusions

Medical science today is too competitive, too big, too entrepreneurial, and too much bent on winning. Gone is the threadbare gentility when science depended on philanthropy for its support and was carried out by a solo investigator with perhaps one fellow and one technician (5). That era is gone if for no other reason than our science today is technically so sophisticated. But while

science may be less genteel, less threadbare, and less fun, none of these is an excuse for fraud or lesser abuses. Those who have chosen science for a career have, in a sense, taken an oath to discover and disseminate the truth, much as physicians have sworn the Oath of Hippocrates. Beyond these individual commitments, the institutions in which science is performed—the medical schools, research institutes, and universities—need to create an environment in which truth prevails. They must have mechanisms (6) in place to ferret out untruths and to set their house in order, when untruths are discovered.

References

1. Barrett PM. The premed machine. *The Washington Monthly* 1985;May:41-51.
2. Petersdorf RG. The evolution of departments of medicine. *N Engl J Med* 1980;**303**:489-96.
3. Petersdorf RG. Is the establishment defensible? *N Engl J Med* 1983;**309**:1053-7.
4. Petersdorf RG. The case against tenure in medical schools. *JAMA* 1984;**251**:920-4.
5. Petersdorf RG. Academic medicine: no longer threadbare or genteel. *N Engl J Med* 1981;**304**:841-3.
6. Petersdorf RG. Preventing and investigating fraud in research [Editorial]. *J Med Educ* 1982;**57**:880-1.

Irresponsible Authorship and Wasteful Publication

EDWARD J. HUTH, M.D.; Philadelphia, Pennsylvania

Fraud is a dramatic offense in scientific publishing but other offenses are more frequent and probably far more damaging. The most frequent lesser offenses are irresponsible authorship and wasteful publication. The authorship problems include listing "authors" who made little or no contribution to the work reported and omitting of persons who made major contributions. Wasteful publication includes dividing the results in a single study into two or more papers ("salami science"); republishing the same material in successive papers (which need not have identical format and content); and blending data from one study with additional data to extract yet another paper that could not make its way on the second set of data alone ("meat extenders"). Wasteful publication may be the most frequent of these offenses and is probably the most damaging because of its economic implications for publishers, readers, libraries, and indexes.

FRAUD in scientific publication is a dramatic abuse and it readily catches the eye of the popular press and the public. But there are abuses that because of their frequency and ubiquity should concern us more than the extreme of fraud. These are abuses of two kinds: irresponsible authorship and wasteful publication.

Public evidence of these abuses is scanty and largely anecdotal, although they have been discussed in the literature (1, 2). Aside from May's study (3) of duplicate papers in one small corner of mathematics, however, I do not know of any thorough and quantitative studies of the extent and frequency of these problems.

Irresponsible Authorship

I take authorship to mean that an author who is the sole author of a paper has been solely responsible for the work it reports and that all coauthors in a multiauthor paper have been sufficiently involved in the reported work and in the writing of the paper to be able to take public responsibility for its contents (4).

The first species of irresponsible authorship is *unjustified authorship*. There are the authors who simply made technical suggestions and did not take part in the research, did not help write the paper, and did not see the final version submitted to the journal. There is the departmental chairman who did not take part in the research. There is the laboratory technician who did routine work that would have been done regardless of the study and who could not defend the paper's content. There is the principal author's spouse listed as second author on a paper about bowel disease who has a degree in library science and gathered numbers from clinical records.

Unjustified authorship is not simply a problem of numbers. Growing numbers of authors per paper may be justified by growing complexity of research, but editors are seeing numbers that seem disproportionately large for the work the papers seem to represent.

Incomplete authorship, the second species of irresponsible authorship, is probably rare. This abuse is a failure to include as authors persons who had responsibility (5) for critical content of a paper.

Wasteful Publication

Abuses of authorship rarely seem to damage the efficiency of science or seriously sap its resources. But they do undermine the ethic of honesty. Abuses that may be doing more damage involve wasting the resources of scientific publication.

The abuse of *divided publication* is the breaking down of findings in a single piece of research into a string of papers, each of which is what Broad (1) has tagged as the "LPU" or "Least Publishable Unit." The research could probably have been reported in a single paper. But why should the investigators confine themselves to one paper when they can slice up data and interpretations into two, three, four, five, or more papers that will better serve their needs when they face promotion or tenure committees? "Salami science" does not always equal baloney, but such divided publication is often an abuse of scientific publication.

When five papers report findings that could have been reported in one, the editorial process, including peer review, has been turned on five times instead of once. Peer review is expensive in time and in money, for journal offices and for reviewers. Divided publication produces unnecessary press and postage costs for journals and the costs of multiple indexing and abstracting for what should have been represented once. Should we squander the resources of science in such multiplied costs?

Repetitive publication—republishing essentially the same content in successive papers or in chapters, reviews, and papers—also wastes resources. In one episode involving this journal about 2 years ago, we published a paper about a serious adverse pulmonary effect due to a new cardiovascular drug. A paper published in a radiological journal five months later described the same effect in apparently the same patients. The authors of the two papers were different but the papers came from the same medical school. Case details strongly suggested that the same cases had been reported twice. At least one author of the second paper must have been aware of the first paper because that author had been acknowledged for reviewing it. An excuse for this duplication might have been to get the message to two audiences unlikely to see each other's journals. But is the cost of duplicating the message for two audiences justified? For some decades the

Annals of Internal Medicine. 1986;104:257-259.

©1986 American College of Physicians

cost for searching through secondary services for information in a journal not usually seen by the searcher might have justified such duplicative publication. Now, the ease, speed, and relatively low cost of searching scientific literature through online bibliographic services discredit this excuse. One good reason to object to the repetition in this episode is that the re-reporting of the four cases misrepresents the incidence of the reported adverse drug effect. A literature search would turn up eight cases instead of four. A closely related abuse is producing the papers I call "meat extenders," papers that add data to the authors' previously reported data, without reaching new conclusions.

None of these abuses is dramatically unethical in the sense that fakery and fraud are. The scientific community might not even agree on whether repetitive and duplicative publication are unethical. Wasteful publication might be seen as justified by needs to compete for institutional and financial support to ensure academic survival. Scientists do not realize that they help pay the excessive costs for peer reviewers, editors, professional societies, bibliographers, and libraries if they subscribe to journals or use any of these services.

Possible Remedies

These abuses arise from the value publishing brings to individuals and particularly to their institutions. Publications are accomplishments that lead to income and power. I am skeptical that enunciating standards for scientific publication, no matter how rational, will readily influence what is in essence an economic problem. Science involves big money and attracts highly competitive career seekers. Short of a basic change in how science is rewarded and rewards itself, what might be done about abuses of authorship and wastes in publication? Dr. Angell proposes in her paper in this section (6) what might be the best remedy; let me suggest some others.

A first major step could be wide and public acceptance among all scientific disciplines of clearly defined and specific criteria for authorship. Very few disciplines appear to have such criteria. The American Psychological Association has developed criteria (7); the Council of Biology Editors has published a short statement, not full criteria, on authorship (8); and the American Chemical Society has published a statement (9) on ethics of publication that covers authorship. The International Committee of Medical Journal Editors has recently issued several short statements defining authorship (10-12).

A second step could be specific requirements on authorship by scientific journals: an acknowledgment or footnote in every co-authored paper could state the exact contribution by each author that has justified authorship. Alternatively, the statement might be required in an accompanying letter. This procedure might reduce the number of abuses by at least forcing application of some kind of criteria.

A third step, one *Annals* took a few years ago, could cut back the abuse of naming persons as authors who have not seen the final version of the paper. We require that each author sign a "Conditions of Publication" form before we publish the paper. Each author thus affirms having seen the final version and agrees to its publication. Such a form could also require authors to indicate their particular contributions to the paper, such as study design, analysis of data, writing the first draft. Even if such required forms did not eliminate all abuses, they might at least raise awareness that authorship implies having taken on responsibilities and not just willingness to accept credit.

Divided publication may be more difficult for editors to deal with. When a piece of work has been chopped up for two or more papers, they are usually submitted to different journals at about the same time. Unless the authors are honest enough to cite the other papers, the peer reviewers and editors are not likely to be aware of the divided pieces. Editors who become aware of the multiple papers may face problems in asking the authors to withdraw one or more papers from the other journals and to combine them.

Editors can probably do more about repetitive publication. Editors can ask peer reviewers to call attention to any prior publication of the content of the paper under review. Many reviewers do this now, some even checking through online bibliographic searchers, but more might do so if asked specifically.

Editors can require authors to affirm by signature on an appropriate form that the present paper or the essence of its content has not been accepted for publication or already published elsewhere. This procedure could raise authors' awareness that editors are alert to such abuses.

Editorial offices could search bibliographic databases to see whether a similar if not identical paper has already been published. A few years ago this procedure would have been too costly in time and money for most journals when the search had to be carried out with paper indexes that often were not current. Now that we have relatively cheap and up-to-date online services, most journals could probably afford to use such searches for papers to be accepted.

All of us in science have ethical obligations to maintain honest authorship and to prevent wasting space in the expensive vehicles that are scientific journals. Our resources are not unlimited and must be shared.

▶ Requests for reprints should be addressed to Edward J. Huth, M.D.; Editor, *Annals of Internal Medicine*, 4200 Pine Street, Philadelphia, PA 19104.

References

1. Broad WJ. The publishing game; getting more for less. *Science*. 1981;211:1137-9.
2. Kronick DA. *The Literature of the Life Sciences: Reading, Writing, Research*. Philadelphia: ISI Press; 1985:67.
3. May KO. Growth and quality of the mathematical literature. *Isis*. 1968;59:363-71.
4. [Huth EJ]. Authorship from the reader's side [Editorial]. *Ann Intern Med*. 1982;97:613-4.
5. Relman AS. Responsibilities of authorship: where does the buck stop? *N Engl J Med*. 1984;310:1048-9.
6. Angell M. Publish or perish: a proposal. *Ann Intern Med*. 1986;104:261-2.
7. *Publication Manual of the American Psychological Association*. Washington, D.C.: American Psychological Association; 1983:20-1.
8. Ethical conduct in authorship and publication. In: CBE Style Manual Committee. *CBE Style Manual: A Guide for Authors, Editors, and Publishers in the Biological Sciences*. 5th ed. Bethesda, Maryland: Council of Biology Editors; 1983:1-6.

9. *American Chemical Society*. ACS ethical guidelines to publication of chemical research. In: *The ACS Style Guide*. Washington, D.C.: American Chemical Society; 1986:217-22.
10. Editorial consensus on authorship and other matters. *Lancet*. 1985;2:595.
11. INTERNATIONAL COMMITTEE OF MEDICAL JOURNAL EDITORS. Guidelines on authorship. *Br Med J*. 1985;291:722.
12. HUTH EJ. Standards on authors' responsibilities [Editorial]. *Ann Intern Med*. 1985;103:797.

Guidelines on Authorship of Medical Papers

EDWARD J. HUTH, M.D.; Philadelphia, Pennsylvania

IN THE belief that authors and potential authors may be helped by explicit statements of justification for authorship, the following guidelines are offered for research papers, case-series analyses, case reports, review articles, and editorials. These guidelines are based on statements issued by the International Committee of Medical Journal Editors (ICMJE) that were published recently (1-3) and will eventually be incorporated into the Committee's document, "Uniform Requirements for Manuscripts Submitted to Biomedical Journals" (4, 5).

The ICMJE statements on authorship represent five principles that are applied below in the specific guidelines. (These principles have been developed directly from the ICMJE statements with some clarifications.)

Principles for Authorship

Principle 1. Each author should have participated sufficiently in the work represented by the article to take public responsibility for the content.

Investigators in the sciences and others who use scientific information must have confidence in its accuracy and validity. Such confidence rests in part on knowing that at least one person has taken public responsibility for the information, and for published information the responsible persons are authors (6, 7). "Public responsibility" means that an author can defend the content of the article, including the data and other evidence and the conclusions based on them. Such ability can come only from having participated closely in the work represented by the article and in preparing the article for publication (7). This responsibility also requires that the author be willing to concede publicly errors of fact or interpretation discovered after publication of the article and to state the reasons for error. In the case of fraud or other kinds of deception attributable to one or more authors, the other authors must be willing to state publicly the nature and extent of deception and to account as far as possible for its occurrence.

Principle 2. Participation must include three steps: (1) conception or design of the work represented by the article, or analysis and interpretation of the data, or both; (2) drafting the article or revising it for critically important content; and (3) final approval of the version to be published.

The work represented by a scientific article includes forming the hypothesis tested by the research it reports or forming the question it answers, developing the means of gathering the reported data or other evidence, collecting the data or the other evidence, critically analyzing the evidence and any counter-evidence, and writing the article so that it reports accurately all of these steps and their products in the structure of critical argument (8). Authors could not publicly defend the intellectual content of an article unless they understand thoroughly the basis for its origin (conception) and can testify to the validity of its argument (critical analysis of evidence). Authors must also have sufficient involvement in writing the paper, either in drafting the initial version or revising subsequent versions to insure validity of the argument and conclusions, to be able to defend the article as an accurate report (7) of the work that led to it.

Principle 3. Participation solely in the collection of data (or other evidence) does not justify authorship.

Data and other evidence bearing on the conclusions and validity of a scientific article may be gathered by persons who know little or nothing of the steps critical to its main intellectual substance: the genesis, design of the work, and the critical analysis of evidence. Such persons could not take public responsibility for the main elements of an article and could testify only to the validity of elements of evidence and not to how they support the argument and its conclusion. Contributions of data by persons for whom authorship is not justified can be acknowledged by other means (*see* Principle 5).

Principle 4. Each part of the content of an article critical to its main conclusions and each step in the work that led to its publication (steps 1, 2, and 3 in Principle 2) must be attributable to at least one author.

"Scientists can proceed with confidence only if they can assume that the previously reported facts on which their work is based are indeed correct" (7). Each element of a scientific article vital to its conclusions therefore must be publicly defensible or its validity is open to question. Therefore, the authorship of a paper must include one or more persons able to defend any of its critically vital elements (9).

Principle 5. Persons who have contributed intellectually to the article but whose contributions do not justify authorship may be named and their contribution described—for example, "advice," "critical review of study proposal," "data collection," "participation in clinical trial." Such persons must have given their permission to be named. Technical help must be acknowledged in a separate paragraph.

Contemporary research can involve persons whose contributions are not vital to the argument and conclu-

▶ From the American College of Physicians, Philadelphia, Pennsylvania.

Annals of Internal Medicine. 1986;104:269-274.

©1986 American College of Physicians

sions of the article but that have been supportive for the authors. Lest authors misrepresent themselves as being solely responsible for all that the article represents, they should indicate who provided intellectual assistance and its nature. Purely technical assistance should be acknowledged separately. Technical assistance includes building equipment, collecting data (specimen gathering and laboratory measurements), locating and abstracting literature, and work in preparing a manuscript that is not intellectual work on its scientific content.

Guidelines for Specific Kinds of Articles

Principle 2 defines participation in three steps as a requirement for authorship:

Step 1. Conception of the work represented by the article, design of the work, analysis and interpretation of data or other evidence presented in the article, or all of these.

Step 2. Drafting the article or revising it for critically important content.

Step 3. Approving the final version of the article for publication.

Steps 2 and 3 necessarily apply to *all* types of articles. Specific contributions in Step 1 may differ for the various types of articles written for a clinical journal.

STEP 1 FOR ARTICLES REPORTING CLINICAL, EPIDEMIOLOGIC, OR LABORATORY RESEARCH

Conception: Framing a specific hypothesis to be tested or specific question to be answered.

Design of the work: Drafting and deciding on the structure and methods for the research.

Analysis and interpretation of the data. This function includes assessing the precision, accuracy, and relevance of data, and statistical analysis. It also includes reviewing the literature for supportive evidence and counter-evidence.

Participation solely in study design or in data analysis or in both may represent adequate participation in Step 1 to justify authorship. Providing technical help, simple referral of patients, or collecting data do not by themselves represent adequate participation in Step 1. In epidemiologic studies the referral of a problem for study does not by itself represent adequate participation in Step 1, but recognizing in the problem an hypothesis to be tested or a specific question to be answered may be adequate.

STEP 1 FOR ARTICLES REPORTING A CASE-SERIES ANALYSIS

Conception: Framing the specific question or questions the analysis is expected to answer.

Design of the work: Defining the characteristics of the cases to be analyzed and the scope of the literature to be considered for supportive evidence and counter-evidence.

Analysis and interpretation of the case data and literature evidence: Critical assessment; structuring and presentation; statistical analysis.

Locating and abstracting case data or literature do not by themselves represent adequate participation in Step 1 for this kind of article. Providing case data ("routine examination and tests") that would have been obtained even if the case-series analysis was not to be carried out is not participation justifying authorship (see Principle 3).

STEP 1 FOR INDIVIDUAL CASE REPORTS

Conception: Recognizing and defining the case characteristics that appear to justify further study of the case and eventually the report.

Design of the work: Deciding on and securing additional case data and relevant literature evidence that support the importance identified in "Conception."

Analysis and interpretation of the case data and literature evidence: Critical assessment and selection of case data and literature evidence. Providing case data (such as "routine tests," laboratory estimations, roentgenographic or other imaging studies, cardiac studies) does not by itself represent adequate participation in Step 1 to justify authorship. Just the referral of the patient (case) to the person or persons responsible for "Conception" does not justify authorship.

STEP 1 FOR REVIEW ARTICLES, EDITORIALS, AND SIMILAR ARTICLES BASED ON CRITICAL ASSESSMENT OF THE LITERATURE AND PERSONAL EXPERIENCE

Conception: Framing the specific question or questions to be answered.

Design of the work: Defining the characteristics of the literature to be reviewed.

Analysis and interpretation of the evidence considered: Selection of evidence through critical assessment.

Locating and abstracting the literature are not by themselves participation in Step 1.

Additional Considerations

TIME OF DECISIONS ON AUTHORSHIP

At least tentative decisions on authorship should be made, if possible, at the beginning of the study, after the potential authors have agreed on its design. Early decisions may reduce the risks of late arguments about authorship. These decisions may have to be revised because of events during the execution of the study, but the final decision should be made no later than the beginning of the first draft of the article (10).

RESPONSIBILITIES FOR DEFINING AUTHORSHIP

Agreement should be reached at the time of the initial decision on authorship as to who will be responsible for subsequent decisions. Generally this person should be the one who will have had the most responsibility for the work in Step 1, defined above. If the first candidate is not available when a final decision on authorship must be reached, the responsibility may be assigned to another person, preferably another author, who is familiar with the conduct of the reported work.

At the time of the initial decision on authorship, agreement should be reached on how subsequent persisting disagreements will be arbitrated. The most impartial arbitration is likely to come from a person familiar with the field of study but not associated with any of the potential authors.

Credit for work done is not a basis for assigning authorship. "Credit" is the appreciation, esteem, or reputation gained by an author after the article has been published, authorship having been assigned before publication on the basis of taking responsibility as discussed above. The credit is assigned to an author by a reader or an institution after estimating the article's importance and whether it represents meritorious work.

SEQUENCE OF AUTHORS

The relative contributions of authors to the intellectually most critical aspects of the work should determine their sequence. Contributions in Step 1 should be given the greatest weight. The first author should have made major contributions in Steps 1 and 2; the following sequence of authors should represent progressively lesser contributions.

THE SINGLE AUTHOR

A single author should take all the responsibilities defined by the five principles set forth above. Single authors must not violate Principle 4, no one who should be taking responsibility for some part of the content of the paper should be omitted from authorship.

CORPORATE (COLLECTIVE) AUTHORSHIP

When a study has been carried out by a group, and no one person can be identified as having a substantially greater responsibility for its content than the others, the group should consider representing themselves by a corporate (collective) title that should reflect the field of the study. The article should carry in a footnote or as acknowledgments the names of the persons represented by the corporate title and of their institutions.

Conclusion

These guidelines are subject to revision in response to needs for additional specific guidelines and to criticism.

Appendix

The guidelines given in the body of this article are largely based on the statements on authorship issued by the International Committee of Medical Journal Editors (1-3); those statements were in turn based in part on previously published opinions and organizational guidelines. I consulted additional documents in preparing the guidelines set forth here. I have selected short quotations from both of these sets of documents to represent their essence and have added brief comments to put them into an historical frame.

1945: WALTER B. CANNON ON AUTHORSHIP

Toward the end of his long and highly fruitful career in physiology, the eminent American physiologist Walter B. Cannon published his scientific autobiography, *The Way of an Investigator* (11).

> The treatment of his collaborators by the head of a laboratory may be put to a real test when the time arrives for publishing papers. Methods which have proved satisfactory in my experience through several decades are as follows. If I have merely suggested the problem to be investigated, indicated the pertinent literature, demonstrated the method to be employed, and from time to time have supervised the work, I have not allowed my own name to appear on the published paper as a joint author. Although the beginner who has carried on research in these circumstances has usually thanked me graciously at the end of his report, that acknowledgement has never been requested. When I have participated in the experimental procedures, my name has appeared as one of the authors. If I have done a major part of the work my name has been placed first, but if my role has been secondary it has not had that prominent position.

1953: MULTIPLE AUTHORSHIP BEGINS TO GROW

The support by foundations of research in the years immediately before World War II and the even greater support by the National Institutes of Health after the war steadily increased the pace of research. As studies became more complex, the number of medical investigators grew rapidly; increasing numbers of claims on authorship apparently were responsible for more frequent multiple authorship (12).

> A major trend in the organization of science during the past two decades has been an evolution away from the individual investigator in favor of research teams. Accompanying this trend, scientific publications have become afflicted with an increasing tendency towards multiple authorship of papers. . . .
>
> At the risk of being presumptuous, we would define the requisites for authorship of a scientific paper as being "the contribution of creative thinking to the advancement of science." Creative thinking assumes . . . its greatest importance in the design of an experiment. How many creative minds contribute to the design of the usual experiment? . . .
>
> Instead of being a means of credit for creative endeavors, it is evident that there is a tendency to degrade authorship into a form of menial patronage. . . . A reversal of present trends will require the stringent elimination of the practice of carelessly offering co-authorship to one's colleagues as a token for small services rendered in the conduct of research.

1957: A NEED FOR ETHICS FOR MEDICAL AUTHORSHIP

Dr. Richard M. Hewitt, senior consultant, Section of Publications, the Mayo Clinic, was thoroughly familiar with medical journals and the problems of writing medical papers, including decisions on authorship. In his classic on medical writing, *The Physician-Writer's Book* (6), he defines a high standard for authorship. Responsibility, not credit, is the essence of Hewitt's standard.

> If we would define publication of unoriginal, repetitious material as not in accord with medical ethics, and would officially reprove it as such, the tawdry author would be silenced and the genuine one helped. A man of the latter type, having something noteworthy to say, may justifiably say it half a dozen times in as many months, but he would prefer to publish it only once. . . . "Johnson was by no means of opinion," wrote Boswell . . . "that every man of a learned profession should consider it incumbent upon him . . . to appear as an author." What would the forthright old man have had to say of the professional man who appeared as an author even though he had not the full right of authorship?
>
> Authorship cannot be conferred; it may be undertaken by one who will shoulder the responsibility that goes with it. To a responsible writer, an article, with his name on it, is the highest product of his mind and art, his property, as nearly flawless as he can make it, founded in his character and evidence of it. If that describes the acceptable standard, medical writers, a responsible group, are in present need of reconsidering the implications of joint authorship. The reader of a report issued by two or more authors has a right to assume that each author has some authoritative knowledge of the subject, that each contributed to the investigation, and that each labored on the report to the extent of weighing every word and quantity in it.

The by-line, then, is not merely a credit line. He who took some part in the investigation, be it ever so minor, is entitled to credit for what he did. If, however, any one of the reader's assumptions mentioned above cannot be applied to such a one, the credit due him, be it ever so major, does not justify his inclusion as an author. His contribution should be acknowledged preferably in the body of the paper, or in a footnote. Further, the generous chap who would bestow authorship on another, perhaps without even submitting the manuscript to him, may do his colleague no favor.

1970: AUTHORSHIP IN PSYCHOLOGY

Other fields have faced problems in authorship. Psychology has had research of growing complexity and thus increasing difficulties with decisions on multiple authorship. Spiegel and Keith-Spiegel (13) saw a consensus in the data from a survey of a large group of psychologists.

. . . most of the psychologists surveyed felt that publication credit should be given only to persons who are very actively involved in contributing to a project and that authorships should never be given out of gratitude or deference to persons of higher status. They should be given when considerable important work has been done, and authorship should reflect the relative significance of the contributions made. Footnotes should be relatively concise indications of minor contributions and should be as explicit as possible concerning the nature of the contribution.

. . . When collaborators cannot agree about authorship order or credit assignment, perhaps some disinterested arbitrators should be engaged to judge the relative merits of the contributions made to the final research product.

1979: COOPERATIVE CLINICAL TRIALS AND AUTHORSHIP

Growth in the number and size of cooperative clinical trials led to questions of authorship for large groups of collaborators. In a 1979 article, Relman gave his definition (14).

Authorship of a paper ought to denote a significant share of responsibility for the conception and design of an experiment and for the analysis and interpretation of the results. At least some personal involvement in the actual gathering of the data is desirable but not always possible. At the very least, however, a coauthor ought to be thoroughly familiar with the experimental details and should feel comfortable about assuming responsibility for them. The essential criterion is the quality of the intellectual input. A scientific paper is a creative achievement, a record of original productivity, and coauthorship ought to be unequivocal evidence of meaningful participation in the creative effort that produced the paper. . . . the use of coauthorship as a kind of payment for faithful technical assistance or data collection violates this principle. So does the all too common practice of adding the chief's name to every paper published from his department or laboratory, regardless of whether he has made any intellectual contribution to the design, execution, or interpretation of results.

1982: RESPONSIBILITY FOR WHAT A PAPER SAYS, NOT ONLY FOR WHAT IT REPRESENTS

A scientific paper is more than a report of research findings and their interpretation. It also represents thought about how to relate its conclusions to the scientific literature and about what to say in print and how. An author is responsible not only for the research reported but for the report itself (15).

An author should have taken part in the writing of the paper.

. . . An author should have read the entire contents of a paper and assented to its publication before it is sent to a journal.

. . . Consider that the person entitled to put his name on a mosaic is not he who gathered the chips but the artist who put together the picture; the two persons may be the same but the claim to creation lies in the picture.

An editorial in *The Lancet* (16) makes a similar point.

Perhaps the least to be expected of authors is that they have participated in and contributed to the published study, that they have read the paper to which they have put their names, and that within the limits of their skills they are prepared to vouch for the work.

1983: STATEMENTS ON AUTHORSHIP FROM PROFESSIONAL SOCIETIES

Scientific organizations with responsibilities for standards in publication are becoming aware that authorship must be defined more precisely if decisions on authorship are to be rational and ethical. In 1983 (17) the American Psychological Association amplified its relevant 1981 ethical principle.

Authorship is reserved for persons who receive primary credit and hold primary responsibility for a published work. Authorship encompasses, therefore, not only those who do the actual writing but also those who have made substantial contributions to a study. This concept is discussed in Principle 7f of the "Ethical Principles of Psychologists" (APA, 1981). . . . *Principle 7f. Publication credit is assigned to those who have contributed to a publication in proportion to their professional contributions. Major contributions of a professional character made by several persons to a common project are recognized by joint authorship, with the individual who made the principal contribution listed first. Minor contributions of a professional character and extensive clerical or similar nonprofessional assistance may be acknowledged in footnotes.*

. . . Substantial professional contributions may include formulating the problem or hypothesis, structuring the experimental design, organizing and conducting the statistical analysis, interpreting the results, or writing a major portion of the paper. Those who do so are listed in the by-line. Lesser contributions, which do not constitute authorship, may be acknowledged in a note. . . . These contributions may include such supportive functions as designing or building the apparatus, suggesting or advising about the statistical analysis, collecting the data, modifying or structuring a computer program, and arranging for research subjects. Combinations of these (and other) tasks, however, may justify authorship. In any case, the writer should always obtain a person's consent before including that person's name in a by-line or in a note.

Authors are responsible for determining authorship and for specifying the order in which two or more authors' names appear in the byline. The general rule is that the name of the principal contributor should appear first, with subsequent names in order or decreasing contribution.

Authors are also responsible for the factual accuracy of their contributions.

A short statement was issued by the Council of Biology Editors (7) in 1983.

The authorship of a paper should be decided when the paper is written, even if the decision is only tentative. This decision should come from the scientist who has been most engaged in designing and executing the research. Any conflicts on authorship or content of the paper should be resolved among the co-workers. The basic requirement for authorship is that an author should be able to take public responsibility for the content of the paper. . . . An author should be able to indicate why and how the observations were

made, and how the conclusions follow from the observations. An author should be able to defend criticisms of the paper, as, for example, in a letter-to-the-editor responding to published criticisms. These abilities should come from having participated in design of the study, in observing and interpreting the reported findings, and in writing the paper.

Claims to authorship may come from persons who have had little to do with the intellectual content of the paper, but who have provided financial support, routine technical assistance, or research space and equipment. Such contributions need not be rewarded with authorship but can be acknowledged in the appropriate section of the paper. . . . A paper must report only observations actually made by one or more of its authors. . . . At least one of the authors, preferably the principal author, should have been closely enough involved with conduct of the study to be reasonably sure that data have not been fabricated or improperly manipulated by any of the other authors or by technicians. . . .

1984: A SUCCINCT STATEMENT OF CRITERIA FOR AUTHORSHIP

Dr. Claude T. Bishop, Editor-in-Chief, National Research Council of Canada Research Journals, has stated (18) succinct criteria.

The authors of scientific papers clearly bear the full responsibility for the veracity of the work reported therein. . . . who warrants coauthorship? . . . Since authorship implies responsibility, one simple guideline could be that all authors should be capable of participating in a discussion or defense of their paper.

1984: AUTHORSHIP CONSIDERED IN THE CONTEXT OF ETHICAL SCIENCE

Sigma Xi, the scientific research society, has issued a broad but detailed discussion (19) of ethical and unethical conduct in science. Several passages cover the questions of legitimate authorship.

The literature of science does have a long tradition of cooperation among two or three authors, who come together because each can contribute specialized knowledge, or because research is often more exciting and rewarding if it is not done entirely alone. What is comparatively new is the practice, in some disciplines, of publishing reports in which five or even fifty individual scientists claim "authorship" of the same paper. It is particularly evident in some forms of biomedical research. . . . If the research requires such cooperation, it is argued that those who contributed should be credited with authorship of the report.

What, it might be asked, has this to do with scientific honesty? How is multiple authorship related to our taxonomy of trimming, cooking, forging and plagiarism? Nothing in principle, perhaps but it seems evident that multiple authorship increases the opportunity for each of these to occur, if only because the responsibilities of authorship are diffused or diminished when they are widely shared.

Irresponsible authorship, rather than multiple authorship, is in fact the real problem in such situations. In principle, it is possible for fifteen or fifty scientists to coauthor a single research report, using the term "author" in the full sense of that word. More usually, however, multiple authorship indicates a claim for credit rather than an acceptance of responsibility. . . . Too often, someone is named as an author less because of the need to accord appropriate recognition than because a publication list is regarded as the index of a scientist's worth, and the more the better.

. . . If [a] paper contains fraudulent statements, or mistakes caused by the carelessness or self-deception of others, it should not have been published and you should not have attached your name and scientific reputation to it as a coauthor. In short, the time to take responsibility for a paper is not after its errors have been exposed but before. Whatever view of the matter is taken by other coauthors, it is up to you to ensure that the manuscript is free of error or bias.

. . . Authorship . . . should mean the same thing in any branch of science.

1985: STANFORD'S PRESIDENT CONSIDERS PROBLEMS OF ACADEMIC AUTHORSHIP

Questions on criteria for authorship and their application in increasingly complex academic settings have been recommended by Dr. Donald Kennedy, President of Stanford University, to his Academic Council as needing systematic discussion. Two paragraphs in his request to the Council (20) focus sharply on the main problems in authorship.

. . . faculty members often rely on their own familiarity with the conventions of the discipline regarding coauthorship and other forms of credit, forgetting that students and other participants in a project may not be acquainted with them. Again, I wonder whether departments or laboratories could ameliorate the bitterness of disappointed expectations by a general discussion, in advance, of the ground rules. The understanding in my laboratory was this: If I had contributed to the idea of the project and had also contributed significantly to the hands-on work, coauthorship was justified; but any coauthor had to have a complete enough grasp of the whole record to defend it effectively in a scientific meeting. This test, of course, is tailored to an experimental science and surely is not the only one applicable even there.

. . . There is a tight coupling between authorship and responsibility. Let us suppose that the name of a faculty member has been included on a paper resulting from the relatively independent experiments by a student or fellow. If the data are then shown to be faulty, or worse, invented, it seems clear to me that the faculty member is responsible. Indeed, as Provost Albert Hastorf's 1984 memorandum on academic fraud pointed out, faculty members are generally responsible for the scholarly conduct of staff and students involved in their research enterprises. When one assumes coauthorship, a still higher duty of certainty prevails. The defense of minimal participation in work done in one's laboratory is generally questionable; surely it is entirely inapplicable when one is coauthor of the disputed work.

1986: GUIDELINES ON AUTHORSHIP FROM THE AMERICAN CHEMICAL SOCIETY

Authorship is one aspect of scientific publication considered in "ACS Ethical Guidelines to Publication of Chemical Research" (21), issued in 1986 by the American Chemical Society.

. . . The co-authors of a paper should be all those persons who have made significant scientific contributions to the work reported and who share responsibility and accountability for the results. Other contributions should be indicated in a footnote or an "Acknowledgements" section. An administrative relationship does not of itself qualify a person for co-authorship (but occasionally it may be appropriate to acknowledge major administrative assistance). Deceased persons who meet the criterion for inclusion as co-authors should be so included, with a footnote reporting date of death. No fictitious name should be listed as an author or co-author. The author who submits a manuscript for publication accepts the responsibility of having included as co-authors all persons appropriate and none inappropriate. The submitting author should have sent each living co-author a draft

copy of the manuscript and have obtained the co-author's assent to co-authorship of it.

▶ Requests for reprints should be addressed to Edward J. Huth, M.D., Editor, *Annals of Internal Medicine*, 4200 Pine Street, Philadelphia, PA 19104.

References

1. INTERNATIONAL COMMITTEE OF MEDICAL JOURNAL EDITORS. Guidelines on authorship. *Br Med J*. 1985;**291**:722.
2. Editorial: consensus on authorship and other matters. *Lancet*. 1985;**2**:595.
3. HUTH EJ. Standards on authors' responsibilities. *Ann Intern Med*. 1985;**103**:797.
4. INTERNATIONAL COMMITTEE OF MEDICAL JOURNAL EDITORS. Uniform requirements for manuscripts submitted to biomedical journals. *Ann Intern Med*. 1982;**96**:766-71.
5. INTERNATIONAL COMMITTEE OF MEDICAL JOURNAL EDITORS. Uniform requirements for manuscripts submitted to biomedical journals. *Br Med J*. 1982;**284**:1766-70.
6. HEWITT RM. *The Physician-Writer's Book: Tricks of the Trade of Medical Writing*. Philadelphia: W. B. Saunders Company; 1957:312-3.
7. Ethical conduct in authorship and publication. In: *CBE Style Manual: A Guide for Authors, Editors, and Publishers in the Biological Sciences*. 5th ed. Bethesda, Maryland: Council of Biology Editors, Inc.; 1983:1-2.
8. HUTH EJ. Critical argument: the basic structure of papers. In: HUTH EJ. *How to Write and Publish Papers in the Medical Sciences*. Philadelphia: ISI Press; 1982:47-9.
9. RELMAN AS. Responsibilities of authorship: where does the buck stop? *N Engl J Med*. 1984;**310**:1048-9.
10. HUTH EJ. Preparing to write. In: HUTH EJ. *How to Write and Publish Papers in the Medical Sciences*. Philadelphia: ISI Press; 1982:37-46.
11. CANNON WB. *The Way of an Investigator*. New York: W.W. Norton; 1945:93.
12. ALEXANDER RS. Trends in authorship. *Circ Res* 1953;**1**:281-3. (Reprinted in *Arch Intern Med*. 1970;**125**:771-2).
13. SPIEGEL D, KEITH-SPIEGEL P. Assignment of publication credits: ethics and practices of psychologists. *Am Psychol*. 1970;**25**:738-47.
14. RELMAN AS. Publication and promotions for the clinical investigator. *Clin Pharmacol Ther*. 1979;**25**:673-6.
15. HUTH EJ. Authorship: from the reader's side. *Ann Intern Med*. 1982;**97**:613-4.
16. Author! *Lancet*. 1982;**2**:1199.
17. AMERICAN PSYCHOLOGICAL ASSOCIATION. Characteristics of authorship and articles. In: *Publication Manual of the American Psychological Association*. Washington, D.C.: American Psychological Association; 1983.
18. BISHOP CT. Ethics. In: BISHOP CT. *How to Edit a Scientific Journal*. Philadelphia: ISI Press; 1984:76-8.
19. JACKSON CI. *Honor in Science*. New Haven: Sigma Xi, The Scientific Research Society; 1984:23-7.
20. KENNEDY D. *On Academic Authorship*. Stanford: Stanford University; 1985.
21. AMERICAN CHEMICAL SOCIETY. ACS ethical guidelines to publication of chemical research. In: *The ACS Style Guide*. Washington, D.C.: American Chemical Society; 1986:217-22.

Abuses and Uses of Authorship

THE FIVE papers in the Perspective section of this issue (1-5) call attention to abuses of authorship and to wasteful publication, two ills afflicting medical journals and perhaps other medical media. Are these ills widespread enough to merit such attention? Although the frequency of abuses and waste is not known, three authors of the Perspective papers who work in medical publishing, Drs. Angell, Bailar, and I, have seen enough of these problems to be convinced that medical academia should begin to get to work on reforms.

Four of the five papers suggest specific reforms and possible remedies. The abuses in presenting evidence in research papers discussed by Dr. Bailar (4) can be reduced if editors select only peer reviewers highly competent to assess papers for scientific validity and if editors and their staffs sharpen their skills in identifying evident flaws in papers. There are also some editorial remedies for duplicative and repetitive publication (3), but editors alone cannot do what needs to be done. The abuses certainly stem in part from "pressure to publish," and Dr. Angell's proposed lid on the number of papers academic candidates for appointment, promotion, or tenure can submit for judgment by committees (5) has great merit.

There are signs that some persons with major responsibilities in the academic world of medicine do not understand the ethical standards editors generally assume are well known. For example, in the apparent fraud represented by the Darsee affair some of the responses from Darsee's coauthors suggested that they felt they could simply disown work to which they had put their names (6). If authorship means taking public responsibility for work on which you have put your name, are authors not obliged to explain to the scientific community what faults led to publication of an invalid paper?

A recent newspaper article reports what may be a more ominous sign of academic irresponsibility. Floyd Abrams (7), an expert on libel, refers to threats of legal action against authors of an unpublished paper said to analyze problems of irresponsible authorship involving the Darsee papers.

> . . . a searching article [by two authors at the National Institutes of Health] addresses questions about Dr. Darsee's collaborators. . . . [the] manuscript has never been published in the two years since it was written. Several dozen distinguished senior scientists were given a preliminary draft. Most praised it. But an avalanche of lawyers' letters, some more than 50 pages long, descended upon [the authors], the N.I.H. and the small scientific journals to which their article had been submitted. Written on behalf of some of Dr. Darsee's co-authors, the letters threatened libel actions if the article was published. . . . One lawyer's letter . . . warned that "publication of the . . . paper by a journal with national circulation would constitute a separate cause of action for libel in all 50 states."

If ethical problems in scientific publication cannot be discussed analytically in scientific journals because academia cannot stand public discussion of responsibilities in scholarship, science is in serious trouble. Indeed, problems in authorship appear to be afflicting more disciplines than the biomedical sciences. Dr. Donald Kennedy, President of Stanford University, has recently asked (8, 9) his

faculty to consider the responsibilities of authorship in the whole range of academic disciplines.

Perhaps, as pointed out by Garfield (10), standards for authorship have not been stated clearly (and loudly) enough by organizations responsible for standards in scientific publication. Steps are being taken toward offering clear standards. The International Committee of Medical Journal Editors (ICMJE) has prepared statements on authorship (11-13) that in their present form or a revised form will eventually be published in a new edition of the ICMJE's *Uniform Requirements* . . . document. Guidelines for specific types of papers appear in the Position Papers section of this issue (14). These guidelines may be copied freely for free distribution by anyone; they are open to revision, and criticisms should be sent to the Editor. The appendix to the guidelines includes excerpts from guidelines of several professional societies, notably the American Psychological Association, the American Chemical Society, and the Council of Biology Editors.

Guidelines alone will not halt or radically reduce current abuses of authorship. It is evident the problems are not going away. The episode discussed by Dr. Petersdorf (1) surfaced just a few months ago in the *Journal of the American College of Cardiology* (15, 16). Academic institutions set standards for scholarship and establish the mores that maintain its integrity. What are medical schools going to do to maintain the reliability and integrity of scientific publishing? The Association of American Medical Colleges (AAMC) responded to episodes of fraud several years ago by issuing guidelines (17). The AAMC should now take a look at the academic setting that is a fertile bed for fraud, in response to Petersdorf's (1) and Woolf's (2) conjectures, and at current and potential criteria for authorship. It should also consider establishing standards on using publications as measures of qualifications for academic appointment, promotion, and tenure.—EDWARD J. HUTH, M.D.; *Editor.*

References

1. PETERSDORF RG. The pathogenesis of fraud in medical science. *Ann Intern Med.* 1986;104:252-4.
2. WOOLF PK. Pressure to publish and fraud in science. *Ann Intern Med.* 1986;104:254-6.
3. HUTH EJ. Irresponsible authorship and wasteful publication. *Ann Intern Med.* 1986;104:257-9.
4. BAILAR J, III. Science, statistics, and deception. *Ann Intern Med.* 1986;104:259-60.
5. ANGELL M. Publish or perish: a proposal. *Ann Intern Med.* 1986;104:261-2.
6. RELMAN AS. Lessons from the Darsee affair. *N Engl J Med.* 1983;308:1415-7.
7. ABRAMS F. Why we should change the libel law. *New York Times Magazine.* 1985,Sept 29: 34, 87, 90, 92, 93.
8. CULLITON BJ. Stanford president calls for new authorship policy. *Science.* 1985;230:422-3.
9. KENNEDY D. *On Academic Authorship.* Stanford, California: Stanford University; 1985.
10. GARFIELD E. More on ethics of scientific publication: abuses of authorship attribution and citation amnesia undermine the reward system of science. *Current Contents.* 1982;25:5-10.
11. INTERNATIONAL COMMITTEE OF MEDICAL JOURNAL EDITORS. Guidelines on authorship. *Br Med J.* 1985;291:722.
12. Editorial consensus on authorship and other matters. *Lancet.* 1985;2:595.
13. HUTH EJ. Standards on authors' responsibilities [Editorial]. *Ann Intern Med.* 1985;103:797-8.
14. HUTH EJ. Guidelines on authorship of medical papers. *Ann Intern Med.* 1986;104:294-7.
15. DACK S. Retraction of two articles. *J Am Coll Cardiol.* 1985;6:1440.
16. DACK S. The continuing problem of retracted papers. *J Am Coll Cardiol.* 1985;7:244-5.
17. ASSOCIATION OF AMERICAN MEDICAL COLLEGES. The maintenance of high ethical standards in the conduct of research. *J Med Educ.* 1942;57:896-904.

©1986 American College of Physicians

APPENDIX 8

PROPOSED NATIONAL SCIENCE FOUNDATION REGULATION FOR DEALING WITH MISCONDUCT IN SCIENCE AND ENGINEERING RESEARCH

4158 Federal Register / Vol. 52, No. 27 / Tuesday, February 10, 1987 / Proposed Rules

NATIONAL SCIENCE FOUNDATION

45 CFR Part 689

Misconduct in Science and Engineering Research

AGENCY: National Science Foundation.

ACTION: Proposed Regulations.

SUMMARY: The National Science Foundation proposes to issue regulations to establish what the NSF and its staff do if they learn of possible misconduct under an NSF award and if they find actual misconduct under an NSF award. Responsibilities of grantee institutions are also set out.

DATE: Comments must be received on or before April 13, 1987.

ADDRESS: National Science Foundation, Office of the General Counsel, Room 501, 1800 G Street, NW., Washington, D.C. 20550.

FOR FURTHER INFORMATION CONTACT: Arthur J. Kusinski or Sukari Smith, (202 357-7439. (This is not a toll-free number.)

SUPPLEMENTARY INFORMATION: Although the National Science Foundation (NSF) has received relatively few allegations of misconduct or fraud occurring in NSF-supported research, or in proposals for the support of research, allegations of this nature are serious enough to warrant establishing formal policies and procedures to handle them. The Foundation believes that grantee institutions bear primary responsibility for preventing and detecting misconduct, and the proposed regulations set forth the role which institutions are expected to take. The proposed regulations, in establishing policies and internal procedures for handling allegations of misconduct, provide for interim administrative actions and appeals procedures.

The proposed regulations are similar to policies and procedures of the Public Health Service (PHS) of the Department of Health and Human Services.

E.O. 12291, Federal Regulation

I have determined that this is not a major rule as defined under section 1(b) of E.O. 12291, Federal Regulation.

Regulatory Flexibility Act

I certify that this regulation will not have a significant economic impact on a substantial number of small entities because it affects primarily the internal procedures of a Federal agency.

List of Subjects in 45 CFR Part 689

Misconduct, Debarment and suspension, Fraud.

Dated: February 3, 1987

National Science Foundation.

Charles H. Herz,

General Counsel.

Accordingly, the National Science Foundation is proposing to add a new Part 689 to Title 45 of the Code of Federal Regulations as follows:

Title 45 Code of Federal Regulations—Public Welfare

PART 689—MISCONDUCT IN SCIENCE AND ENGINEERING RESEARCH

Sec.

689.1 General policies and responsibilities.
689.2 Actions.
689.3 Role of awardee institutions.
689.4 Initial NSF handling of misconduct matters.
689.5 Investigations.
689.6 Pending proposals and awards.
689.7 Interim administrative actions.
689.8 Dispositions.
689.9 Appeals.

Authority: Sec. 11(a) of the National Science Foundation Act of 1950, as amended [42 U.S.C. 1870(a)].

§ 689.1 General policies and responsibilities.

(a) "Misconduct" means (1) fabrication, falsification, plagiarism, or other serious deviation from accepted practices in proposing, carrying out, or reporting results from research; (2) material failure to comply with Federal requirements for protection of researchers, human subjects, or the public or for ensuring the welfare of laboratory animals; or (3) failure to meet other material legal requirements governing research.

(b) The NSF will take appropriate action against individuals or institutions upon a determination that misconduct has occurred under an NSF award. It may also take interim action during an investigation. Possible actions are described in section 689.2.

(c) NSF will find misconduct only after careful inquiry and investigation by an awardee institution, by another

Federal agency, or by NSF. An "inquiry" consists of information-gathering and preliminary fact-finding to determine whether an allegation or apparent instance of misconduct warrants an investigation. An "investigation" is a formal examination and evaluation of relevant facts to determine whether misconduct has taken place or, if misconduct has already been confirmed, to assess its extent and consequences or determine appropriate NSF action.

(d) Before NSF makes any final finding of misconduct or takes any final action on such a finding, the individual or institution accused of misconduct normally will be given notice, a chance to provide comments and rebuttal, and a chance to appeal.

(e) Debarment, suspension, or termination of an award for misconduct will be imposed only after further procedures described in applicable debarment and suspension regulations.

(f) The Division of Audit and Oversight (DAO) in the Office of Budget, Audit, and Control, oversees and coordinates NSF activities related to misconduct, conducts any NSF inquiries and investigations into suspected or alleged misconduct, and except where otherwise provided, speaks and acts for NSF with affected individuals and institutions. The Office of the General Counsel (OGC) advises DAO and represents NSF on any current or potential criminal prosecution, current or potential litigation, or significant legal questions that arise.

§ 689.2 Actions.

(a) Possible final actions listed below for guidance range from minimal restrictions (Group I) to the most severe and restrictive (Group III). They are not exhaustive and do not include possible criminal sanctions.

(1) *Group I actions.* (i) Send a letter of reprimand to the individual or institution.

(ii) Require as a condition of an award that for a specified period an individual, department, or institution obtain special prior approval of particular activities from NSF.

(iii) Require for a specified period that an institutional official other than those guilty of misconduct certify the accuracy of reports generated under an award or provide assurance of compliance with particular policies, regulations, guidelines, or special terms and conditions.

(2) *Group II actions.* (i) Restrict for a specified period designated activities or expenditures under an active award.

(ii) Require for a specified period special reviews of all requests for funding from an affected individual, department, or institution to ensure that steps have been taken to prevent repetition of the misconduct.

(3) *Group III actions.* I(i) mmediately suspend or terminate an active award under appropriate NSF regulations.

(ii) Debar or suspend an individual, department, or institution from participation in NSF programs for a specified period after further proceedings under applicable regulations.

(iii) Prohibit participation of an individual as an NSF reviewer, advisor, or consultant for a specified period.

(b) In deciding what actions are appropriate when misconduct is found, NSF officials should consider:

(1) How serious the misconduct was;

(2) Whether it was deliberate or merely careless;

(3) Whether it was an isolated event or part of a pattern;

(4) Whether it is relevant only to certain funding requests or awards or to all requests or awards involving an institution or individual found guilty of misconduct.

(c) Interim actions may include, but are not limited to:

(1) Totally or partially suspending an existing award;

(2) Totally or partially suspending eligibility for NSF awards in accordance with debarment-and-suspension regulations;

(3) Proscribing or restricting particular research activities, as, for example, to protect human or animal subjects;

(4) Requiring special certifications, assurances, or other administrative arrangements to ensure compliance with applicable regulations or terms of the award;

(5) Requiring more prior approvals by NSF;

(6) Deferring funding action on continuing grant increments;

(7) Deferring a pending award;

(8) Restricting or suspending use of individuals as NSF reviewers, advisors, or consultants.

§ 689.3 Role of awardee institutions.

(a) Awardee institutions bear primary responsibility for prevention and detection of misconduct. In most instances, NSF will rely on awardee institutions to promptly:

(1) Initiate an inquiry into any suspected or alleged misconduct;

(2) Conduct a subsequent investigation, if warranted; and

(3) Take action necessary to ensure the integrity of research, the rights and interests of research subjects and the public, and the observance of legal requirements or responsibilities.

(b) If an institution wishes NSF to defer independent inquiry or investigation, NSF expects it to:

(1) Inform NSF immediately if an initial inquiry supports a formal investigation.

(2) Keep NSF informed during such an investigation.

(3) Notify NSF even before deciding to initiate an investigation or as required during an investigation (i) if the seriousness of apparent misconduct warrants; (ii) if immediate health hazards are involved; (iii) if NSF's resources, reputation, or other interests need protecting; (iv) if Federal action may be needed to protect the interests of a subject of the investigation or of others potentially affected; or (v) if the scientific community or the public should be informed.

(4) Provide NSF with the final report from any investigation.

(c) If an institution wishes NSF to defer independent inquiry or investigation, it should complete any inquiry and decide whether an investigation is warranted within 30 days. It should similarly complete any investigation and reach a disposition within 120 days. If completion of an inquiry or investigation is delayed, but the institution wishes NSF deferral to continue, NSF may require submission of periodic status reports.

(d) Awardee institutions should maintain and effectively communicate to their staffs appropriate policies and procedures relating to misconduct, which should indicate when NSF must or should be notified.

§ 689.4 Initial NSF handling of misconduct matters.

(a) NSF staff who learn of alleged misconduct will promptly and discreetly inform DAO or refer informants to DAO.

(b) To the extent possible the identity of informants who wish to remain anonymous will be kept confidential.

(c) If alleged misconduct may involve a crime, DAO will promptly consult with OGC, which will determine whether any criminal investigation is already pending or projected. If not, OGC and DAO will determine whether the matter should be referred to the Department of Justice.

(d) Otherwise DAO may:

(1) Inform the awardee institution of the alleged misconduct and encourage it to undertake an inquiry;

(2) Defer to inquiries or investigations of the awardee institution or of another Federal agency;

(3) At any time proceed with its own inquiry.

(e) If DAO proceeds with its own inquiry it will normally complete the

inquiry no more than 60 days after initiating it.

(f) On the basis of what it learns from an inquiry and in consultation as appropriate with other NSF offices, DAO will decide whether a formal NSF investigation is warranted.

§ 689.5 Investigations.

(a) When an awardee institution or another Federal agency has promptly initiated its own investigation, DAO may defer any NSF inquiry or investigation until it receives the results of that external investigation. If it does not receive the results within 120 days, DAO will ordinarily proceed with its own investigation.

(b) If DAO decides to initiate an NSF investigation, it must give prompt written notice to the individuals or institutions to be investigated, unless notice would prejudice the investigation or unless a criminal investigation is under way or under active consideration. If notice is delayed, it must be given as soon as it will no longer prejudice the investigation or contravene requirements of law or Federal law-enforcement policies.

(c) If a criminal investigation by the Department of Justice, the Federal Bureau of Investigation, or another Federal agency is under way or under active consideration by these agencies or the NSF, OGC will advise DAO what information, if any, may be disclosed to the subject of the investigation or to other NSF employees.

(d) An NSF investigation *may* include:

(1) Review of award files, reports, and other documents already readily available at NSF or in the public domain;

(2) Review of procedures or methods and inspection of laboratories, laboratory materials, specimens, and records at awardee institutions;

(3) Interviews with parties or witnesses;

(4) Review of any documents or other evidence provided by or properly obtainable from parties, witnesses, or other sources;

(5) Cooperation with other Federal agencies;

(6) Opportunity for the subject of the investigation to be heard; and

(7) Full adjudicatory hearings or other formal proceedings, as described in appropriate regulations.

(e) NSF may invite outside consultants or experts to participate in an NSF investigation. They should be appointed in a manner that ensures the official nature of their involvement and provides them with legal protections available to federal employees.

(f) DAO will make every reasonable effort to complete an NSF investigation and to report within 120 days after initiating it. If DAO cannot report within 120 days, it should submit to the Deputy Director within 90 days an interim report and an estimated schedule for completion of the final report.

§ 689.6 Pending proposals and awards.

(a) Upon learning of alleged misconduct DAO will identify potentially implicated awards or proposals and, when appropriate, will ensure that program and DGC officials handling them are informed (subject to § 689.5(c)).

(b) Neither a suspicion or allegation of misconduct nor a pending inquiry or investigation will normally delay review of proposals. To avoid influencing reviews, reviewers or panelists will *not* be informed of allegations or of ongoing inquiries or investigations. However, if allegations, inquiries, or investigations have been rumored or publicized, the responsible Assistant Director may, in consultation with DAO, either defer review or inform reviews of the status of the matter.

§ 689.7 Interim administrative actions.

(a) After an inquiry or during an external or NSF investigation the Deputy Director may order that interim actions (as described in § 689.2(c)) be taken to protect Federal resources or to guard against continuation of any suspected or alleged misconduct. Such an order will normally be issued on recommendation from DAO and in consultation with DGC, OGC, the responsible Directorate, and other parts of the Foundation as appropriate.

(b) Such interim actions may be taken whenever information developed during an investigation indicates a need to do so. Any interim action will be reviewed periodically during and investigation and modified as warranted.

(c) The Deputy Director will make and DAO will retain a record of interim actions taken and the reasons for taking them.

(d) Interim administrative actions are not final agency actions subject to appeal.

§ 689.8 Dispositions.

(a) After receiving a report from an external investigation by an awardee institution or another Federal agency DAO will assess the accuracy and completeness of the report and whether the investigating entity followed usual and reasonable procedures. It will either recommend adoption of the findings in whole or in part or, normally within 30 days, initiate a new investigation.

(b) When any satisfactory external investigation or an NSF investigation fails to confirm alleged misconduct and the Deptuy Director concurs,

(1) DAO will notify the subject of the investigation and, if appropriate, those who reported the suspected or alleged misconduct. This notification may include the investigation report.

(2) Any interim administrative restrictions that were imposed will be lifted.

(c) When any satisfactory investigation confirms misconduct,

(1) Except in unusual circumstances, the investigation report will be provided by DAO to the subject of the investigation, who will be invited to submit comments or rebuttal. Comments or rebuttal submitted within the period allowed, normally thirty days, will receive full consideration and may lead to revision of the report or of a recommended disposition.

(2) Normally within 45 days after completing an NSF investigation or receiving the report from a satisfactory external investigation, DAO will submit to the Deputy Director the investigation report, any comments or rebuttal from the subject of the investigation, and a recommended disposition. The recommended disposition will propose any final actions to be taken by NSF. Section 689.2 lists possible final actions and considerations to be used in determining them.

(d) The Deputy Director will review the investigative report and DAO's recommended disposition. Before issuing a disposition the Deputy Director may initiate further hearings or investigation. Normally within thirty days after receiving DAO's recommendations or after completion of any further proceedings, the Deputy Director will send the affected individual or institution a written disposition, specifying actions to be taken. The decision will include instructions on how to pursue an appeal.

§ 689.9 Appeals.

(a) In case of debarment, suspension, or termination of an award for misconduct, the appeals provided for in NSF regulations will be available. In all other cases, an affected individual or institution may appeal to the director in writing within 30 days after receiving the Deputy Director's written decision. The Deputy Director's decision becomes a final administrative action if it is not appealed within the 30 day period.

(b) The Director may appoint an uninvolved NSF officer or employee to review an appeal and make recommendations.

(c) The Director will inform the appellant of a final decision within 30 days after receiving the appeal. That decision will be the final administrative action of the Foundation. Findings from completed investigations may be shared with scientific review groups if the information bears directly on an investigator's scientific intergrity or if necessary to provide an accurate account of relevant facts.

[FR Doc. 87-2705 Filed 2-9-87; 8:45 am]

BILLING CODE 7555-01-M

APPENDIX 9

ADDITIONAL MATERIALS FOR THE RECORD

Report of the Association of American Universities Committee on the Integrity of Research

Approved by
The Joint Committee on Health Policy
of
The Association of American Universities
The American Council on Education
and
The National Association of State Universities
and Land-Grant Colleges

Association of American Universities Committee on the Integrity of Research

Chancellor William H. Danforth, Chairman
Washington University

Dr. Struther Arnott
Vice President for Research
and Dean of the Graduate School
Purdue University

Dr. Albert A. Barber
Vice Chancellor—Research
University of California, Los Angeles

Dr. James Anthony Ferrendelli
Professor of Neurology and Pharmacology
Washington University
School of Medicine

Dr. Richard S. Ross
Dean of the Medical Faculty
The Johns Hopkins University

Dr. Howard G. Sachs
Associate Provost for Academic Affairs
Administration
Ohio State University

Mr. Cliff K. Travis
University Counsel
Indiana University

Introduction

Examples of fraud in research have recently been reported in great detail by the media; these reports have raised concern in the public as well as among the scientists. Many concerned citizens have wondered whether such activity has become more prevalent, whether it is being reported more often, whether it can be detected readily, whether institutions are organized appropriately to inquire into alleged frauds, and whether procedures for administering discipline are in place.

On October 20, 1981, the Association of American Universities voted to establish a Working Group on the Integrity of Research. The resolution read in part:

> "Incidents of misconduct which raise concern about integrity in scientific research have come to our attention. Although we believe such instances to be rare for such a large enterprise as university research, even rare occurrences are unacceptable.
>
> "The AAU therefore recognizes a need for universities to collaborate with professional societies and related organizations in the examination of the sources of such problems and remedies available to them."

The working group subsequently appointed by the Executive Committee of the AAU believes that the integrity of the research process is an essential part of our intellectual and social structure and must be maintained at all costs. Although serious violations of honesty in research may be rare, those that do occur strike at the very heart of the enterprise. Advances in knowledge depend on trustworthy data and honestly reported conclusions. Anything less will seriously undermine the total enterprise and erode public confidence in those responsible for its conduct.

The integrity of the research process must depend largely on self-regulation; it is the responsibility of all who engage in the search for knowledge. This principle has served science in an exemplary way for centuries. Advances are gleaned from rigorous application of scientific methods and in compliance with ethical codes rooted in intellectual honesty.

Deviations from the norm—even serious ones—have usually been dealt with informally and quietly. Although these methods may have generally worked well in the past, experience suggests that it is now appropriate to give serious thought to better methods for preventing and detecting irregularities and to the manner in which universities deal with them.

This committee has looked at some of the questions especially important to the operation of academic institutions. We encourage others such as scientific societies, editors, and funding agencies to look at the matter from their own perspectives.

Types of Fraud or Deviance in Academic Research

Deviant actions by researchers may be grouped in four categories—scholarly fraud by falsification of data, plagiarism, abuse of confidentiality, and deliberate violations of regulations.

Falsification of data undermines the basic principle on which the scientific process depends. Since scientific advances depend on accurate collection, analysis, and reporting of information, dishonest reporting misleads others and results in the waste of resources, both human and monetary. If practiced in clinical research, falsification could even be directly dangerous to humans. Falsification of data ranges from sheer fabrication through selective reporting, including the omission of conflicting data.

Plagiarism is especially hurtful to individual researchers since it is an attempt by one individual to receive credit for the work of someone else. Outright plagiarism is generally easily detected in areas of research that are very actively pursued and is, therefore, rare. However, the academic community tolerates more than it should, more subtle deviations from the ideal. Inadequate citation and parsimony in referencing submission of the same data in more than one publication by the same author, and similar abuses, do occur with some regularity.

Abuse of confidentiality is a significant act of fraud in an environment that depends on peer review. It is quite distinct from plagiarism and more difficult to detect since such abuse does not usually involve verbatim duplication of another's work. In the present environment, researchers freely discuss their ideas in research proposals submitted to potential sponsors. Proposals usually include extensive data to support the ideas. The ideas and preliminary data may be reviewed by departmental colleagues, university committees and administrators, as well as extramural professional peers serving on review panels. In addition,

detailed studies are submitted to professional journals and subjected to further review by professional colleagues long in advance of eventual publication. Opportunities to abuse confidentiality arise at many points during these processes. Moreover, abuse of confidentiality can occur not only by the actions of the primary reviewers but also by the actions of those with whom the reviewers have shaged the privileged information. In many ways confidentiality is the easiest research ethic to abuse and the most difficult to detect.

Instances of *seemingly deliberate violations of regulations* applicable to research have also become a recent problem. Serious violations, especially of rules adopted by appropriate mechanisms to protect patients, research subjects, other persons, and animals, while not fraudulent in the traditional sense, must be considered so deviant as to undermine the integrity of the research process.

Prevention of Dishonesty in Scientific Research

The rewards associated with success tempt certain individuals into dishonest behavior. Scientific discoveries are rewarded by recognition by peers and, if sufficiently interesting or important, by the general community. In addition, productivity reaps tangible rewards, including career advancement, increase in salary, promotion, election to academic societies, receipt of prizes, funds for additional research, and other benefits. Thus, it is not surprising that dishonesty could emerge in some individuals who are more interested in the rewards than in acquiring new knowledge.

Identification in advance of those susceptible to dishonest behavior is desirable. Careful examination of the record of a prospective investigator can prove helpful. Special attention should be given to motivation and integrity at times of recruitment and advancement in responsibility. Credentials and claimed accomplishments should be examined carefully. Scholastic ability and technical competence do not necessarily indicate that the aptitude for science exists. It is often advisable to look beyond the most recent employment or educational experience for evidence of scientific aptitude and capability.

Since dishonesty is an unfortunate response to environmental temptations and since it is difficult if not impossible to detect in advance those most susceptible, major attention should be directed toward establishing the best environment. To accomplish this goal, attention should be given to the following issues.

Encouragement of Intellectual Honesty Nothing can substitute for a pervasive attitude of intellectual honesty in the laboratory environment. A recommitment to the ethical standards of science by all its practitioners is absolutely essential. At a minimum these standards include: open communication, submission of work for peer review, avoidance of conflict of interest, and commitment to self-regulation. The encouragement of intellectual honesty is not the responsibility of a few but must be accepted by all persons in the university. Especially, the scientific leaders must set an example for all by assiduously complying with standards of intellectual honesty and must assume the responsibilities associated with the role of mentor. By maintaining high standards, scientific leaders create a climate which discourages dishonesty and fosters unquestionable integrity. It is our opinion that a positive attitude of intellectual honesty does more to prevent dishonesty than any other single factor.

Discouragement of "Success at Any Cost" Obviously this issue is a difficult one. It is impossible to eliminate productivity and success as determinants for promotion and recognition. To do so would discourage achievement and ambition and would probably markedly attenuate research activities. However, the emphasis on quality rather than quantity of research—especially publication—is strongly recommended. It should be recognized that pressure for more publications may not be explicit, but hidden pressures for more frequent reports and papers will prevail if responsible individuals are mute on the subject. An active and frequently expressed attitude stressing quality rather than quantity is necessary.

Acceptance of Responsibility by the Laboratory Director Although everyone involved in science must be active in the prevention of dishonesty, the director of a laboratory who is mentor or supervisor of research must assume special responsibilities. Personnel must receive appropriate supervision and students must be directed by experienced scientists. The director should supervise, teach, and encourage in-depth scrutiny and interpretation of results, emphasizing respect for primary data. Routine audit and review of all primary data by the laboratory director is strongly recommended. It is inadvisable for the director to delegate these important functions.

The director must assume absolute responsibility for the validity of all communicated information from his or her laboratory and should encourage the publication of as much primary data as possible.

Maintenance of Professional Interpersonal Relationships Interactions among laboratory personnel are important in determining attitudes concerning honesty and dishonesty. Laboratory directors should encourage investigators to work with other colleagues, to share data, and to discuss results freely. Secrecy about methods and data should be discouraged. Directors should also promote a close but open and professional interaction among investigators and between faculty and students. A sense of competition among laboratory personnel or between students and faculty must be avoided. Relationships should be sufficiently personal to encourage openness and freedom of expression but not so close as to interfere with objectivity.

Establishment of Well-Defined Experimental Protocol Well-designed and strictly-adhered-to experimental methods are important deterrents to dishonesty. Written, detailed, explicit procedures for data gathering, storage, and analysis are essential and should be available and practiced in all laboratories. Research which is blinded or coded and the repetition of experiments in the same or a different laboratory should be encouraged.

Appropriate Assignment of Credit and Responsibility A climate of integrity should include generosity in recognizing the accomplishments of others. Adequate citation of the contributions of persons from other laboratories is especially important. Publications should list as authors only those who contributed significantly to the research, are prepared to stand behind the conclusions, and have reviewed the manuscript carefully.

Institutional Policies and Procedures

The committee recommends that all institutions prepare policies which state clearly the expectations for high standards of ethical behavior of those involved in research, the procedures for dealing with suspected deviations from intellectual honesty, and available sanctions. These policies and procedures must be consistent with the institution's policies on academic governance, freedom, responsibility, and due process, as well as with legal restraints.

The committee recommends that the adoption of such policies be given prompt attention by appropriate academic bodies in every university. Some institutions, after consultation with proper faculty committees, might find it suitable to adopt interim procedures with the understanding that a final document will be developed after further consideration.

This committee recommends that the policy statement deal with the following issues.

The Professional Responsibility of Researchers The policy should be explicit about the institutional standards for those engaged in research. In addition, the duties of those with oversight responsibilities should be clear. Mechanisms for periodic review of research policies may be included as necessary.

Procedures for Dealing with Deviations Institutions should have workable procedures for dealing with suspected deviations from intellectual honesty and the authority to apply appropriate sanctions when deviations are proved to the satisfaction of the appropriate body. (The excellent report, "The Maintenance of High Ethical Standards in the Conduct of Scientific Research," adopted by the Association of American Medical Colleges includes a helpful model for dealing with suspected fraud in research.)

Administrative Responsibility An officer or officers of the institution should be designated to inform investigators of policies which affect the conduct of research and to receive and pursue complaints concerning lack of integrity in research. Investigations should not be in the hands of associates from the laboratory in question, since personal relations may make objectivity difficult or impossible. The designated individual or individuals should see to it that appropriate institutional policies are followed and that adequate records are kept. When appropriate, an individual or a committee from within or outside the institution may be appointed to conduct

an investigation. Anyone appointed to investigate suspected fraud must be objective and must possess the special competencies necessary to understand the research in question.

Reporting of Suspected Fraud Members of the academic community have a responsibility to report what they believe to be lack of integrity in research. Policies should provide assurance that such reports will be held in confidence to the extent possible. Persons giving information in good faith about questionable conduct should be protected against reprisals.

The Rights of the Individual A researcher under suspicion should be treated as a colleague whose cooperation in providing access to data and procedures is expected. The individual in question should have ample opportunity to communicate with the investigator or the investigating committee in the course of the inquiry and prior to the formulation of conclusions. The individual should be advised of any decision to disseminate information about the investigation or to seek information about the research from others.

Confidentiality The mere suspicion of wrongdoing, even if totally unjustified, is potentially damaging to an investigator's career. Confidential handling of information about an investigation must be the responsibility of all involved. Thus, information concerning any investigation should be available only to those who need to know. Ideally, an inquiry should remain totally confidential until the results are established with reasonable certainty. Indeed, if the investigation were to conclude that no wrongdoing occurred, the suspicion should be obliterated from memory. However, this ideal is difficult or impossible to attain. This situation may be made easier by recognizing that research methods and results should always be open to inspection, evaluation, and criticism. In this spirit, all involved should be encouraged to accept an investigation of alleged misconduct as part of the process of the search for truth.

External Contacts There might be good reason to inform or consult with external organizations, such as granting agencies, other institutions linked to the research in question, scientific journals, professional disciplinary societies, state licensing authorities and, in some rare instances, police authorities. Such steps should be considered with great care and with appropriate consultation. The university has an obligation to notify a granting agency when the responsible officer concludes that there is a substantial reason to believe that fraud has occurred, even if the investigation has not been completed. This sharing of information would be facilitated if major granting agencies would establish procedures for receiving such information and for sharing in the responsibility to maintain maximum confidentiality until the completion of the investigation. Publicity should be avoided until the investigation concludes.

Use of Facilities and Equipment In some instances the institution might feel compelled to restrict or forbid the accused researcher's use of its premises, equipment, and resources. An institution should not limit or stop research in progress unless continued access to facilities by the alleged wrongdoer is deemed by knowledgeable and informed colleagues to pose a danger to the safety of persons or property or to preclude fair and objective evaluation. Monitored continuation of research, pending resolution of the inquiry, might be considered.

Summary

The AAU-appointed Working Group on the Integrity of Research believes that violations of honesty in research strike at the heart of the scientific enterprise. It is recommended that special attention be given to factors that foster a climate that encourages intellectual honesty. The directors of laboratories have the major responsibility for establishing appropriate standards, but all should share in this effort. It is recommended that each institution develop appropriate policies which include expectations for high standards of ethical behavior, procedures for dealing with suspected deviations, and appropriate sanctions for use when necessary.

On Academic Authorship

Donald Kennedy
President
Stanford University
September 1985

A systematic discussion of two related issues: first, the allocation of responsibility and credit for scholarly work; and second, those forces that are pushing us toward a level of complexity in the conduct of research at which it becomes difficult to determine responsibility of authorship.

September 18, 1985

TO: Members of the Academic Council
FROM: Donald Kennedy
SUBJECT: Academic Authorship

Dear Friends:

For some time, I have felt a need for systematic discussion within this faculty of two related issues: first, the allocation of responsibility and credit for scholarly work; and second, those forces that, in many disciplines, are pushing us toward a level of complexity in the conduct of research at which it becomes difficult to determine responsibility of authorship.

I am sure you are aware of the national attention given to cases involving misconduct and alleged misconduct in the production and publication of research results. You may not know that at Stanford we have been hearing increasing numbers of complaints and disagreements between students (postdoctoral, graduate, and undergraduate) and their faculty sponsors over credit for work to which both may have contributed.

The pattern of research in many fields is changing. Large laboratories may be under the general directorship of one or two individuals but often include relatively independent work being performed by groups of shifting composition. These changes often produce exciting developments, but at the same time they pose novel problems for the allocation of credit and responsibility.

Increased administrative burdens on principal investigators, arising in part out of more onerous government regulation, may also contribute — by decreasing the time available for active participation in research, just when its importance to academic careers is increasing.

For these and perhaps other reasons, matters of authorship, attribution, and acknowledgment have become more complex; responsibility for work in which we are less personally involved has become more common; and the customs prevailing in different fields have diverged. It is also likely that the incentive for claiming credit has increased.

In the past, appropriate standards in these matters have been determined by the individuals involved, based upon traditions that have grown up in the profession as a whole and in the separate fields. But the steady stream of problems now coming before me suggests that isolated individual judgments by faculty members may no longer be adequate. So I write this memorandum in the hope that it may initiate discussion among my colleagues.

Let me first describe how these problems tend to arise. Some actions, of course, remain relatively simple to label — and to condemn. One who has had no connection with a research project but presents its methods, data, or conclusions as his or her own has committed plagiarism; if the writings of another are taken essentially verbatim, there may be copyright infringement as well. I do not think our common understanding about the straightforward theft of intellectual product has been eroded.

But there are more difficult cases, and they generally involve the allocation of credit for work to which several individuals have contributed something. Often, one participant claims that the joint effort permits publication without the consent of the others, or publication as a sole author, or publication without acknowledgment of the contribution of others.

The following scenario is not atypical: Graduate student S is working in Professor P's laboratory on an experimental problem within P's general field of interest and competence. Perhaps S has a research assistantship funded by P's grant. S conducts an experiment, gathers a set of interesting data, and writes up the results as part of what is expected to become a doctoral dissertation. S discusses the results with Professor P, who subsequently incorporates the data and new findings in a symposium article (or perhaps a grant application), with a cursory

acknowledgment of S's role, or none. S considers that original data from an independent experiment was wrongfully appropriated by P. P views the work as having been done with his support, in his laboratory, within the conceptual framework devised by him, and accordingly views the results as a joint intellectual product in which both have full rights of authorship.

Cases of this kind are especially resistant to easy solution because we are such a diverse and complex scholarly community. Each of us will think of factors not given in this bare sketch that would incline us toward one view or the other — but they will not always be the same factors.

Notions about how to handle joint student and faculty work in laboratories in organic chemistry at Stanford and at MIT will resemble one another much more closely than do those applying to organic chemistry and, say, economics within either institution.

In some disciplines it is quite customary for graduate students to publish their own research results by themselves, even when their work involves fairly close supervision by a faculty member — and in others, the professor's name goes on virtually every paper produced in the laboratory.

This heterogeneity of custom sometimes makes it very difficult to decide, in disagreements like the one between S and P, whether we are dealing with professional misconduct (the wrongful appropriation of another's intellectual product) or whether we are in a domain of ethical judgments about the proper allocation of credit between joint researchers — judgments so close that they should be resolved by personal values, etiquette, and generosity, rather than by a faculty disciplinary process. It is especially difficult for the disputing parties to see a matter in the same way.

The resulting level of bitterness can be very high, and the University should bend every effort toward its reduction. Although we probably cannot hope for a set of specific rules that could set criteria for the shared ownership of intellectual product in all scholarly disciplines, some clarification ought to be possible. Here are some propositions that I think are generally true:

First: In settings like ours, research and training are closely related, deeply interpenetrating functions. The research apprenticeship undertaken by any student — undergraduate, graduate, or postdoctoral — in the academic domain of a faculty sponsor has as one of its purposes the training and development of the student. The other purpose, of course, is the advancement of the scholarly program of which the faculty member is the leader. The relationship between the two is complementary; but the functions of the student participation, and its value to the student on the one hand and the faculty member on the other, will vary.

In particular, the functions served and the kinds of benefits to each party change with the student's academic level. In some fields, for instance, undergraduates and beginning graduate students rarely contribute significantly to the intellectual design and content of advanced research; they can be useful as "hands." The benefits they realize may chiefly be that of becoming familiar with fairly routine procedures, learning the methodology of research, the operation of any relevant equipment, and absorbing the "culture" of the discipline. These are substantial benefits, and the contributions they make are helpful to the enterprise without being essentially scholarly. However, we also should be cognizant of our primary responsibility to our students and alert to opportunities to involve them as early as possible in the genuinely intellectual aspects of the discipline. If we keep using them for routine tasks long after they have gained command over them — or if we fail to recognize by appropriate credit in publications that they have progressed and are making intellectual contributions — the relationship becomes exploitative. And even in the case of undergraduates or beginning graduate students, there will be situations in which they are in fact contributing to the intellectual content of the project, and thus functioning as colleagues. To fail to maximize all these opportunities for our students is to default on one of our central missions as educators.

Second: There is a cluster of questions about authorship and intellectual "ownership" that includes these, among others: What level of contribution by the various parties to a research enterprise qualifies for (co)authorship of the product? What circumstances entitle one to independent or to first publication or to the use of data in another publication or project without attribution? In considering these, I am struck by the seamlessness that often characterizes collaborative research, and I wonder if we would do better if we gave thought — and voice — to our assumptions at the outset, even as we understand that those initial assumptions may well change as the course of research progresses.

Even at its beginning, it is often difficult to trace the source of the ideas and insights from which a research project originates; it is hard to say what was identifiably generated by one individual and what was "in the air" in the intellectual domain. Although my personal experience is with the natural sciences, I know that ideas in the social sciences and the humanities are likely to be of similarly elusive, and mixed, parentage. In those sciences where the gathering of experimental data is a regular part of the work, there are matters of experimental concept, design, instrumentation, hands-on execution, and data interpretation; in other fields, an idea is shaped and developed as it is committed to writing, so that the original inspiration may defy reconstruction by the time the project is completed. One member of an experimental team that includes a professor, graduate students, postdoctoral fellows, and technicians may provide an absolutely critical skill, without which the entire venture could not proceed; one contributor to a scholarly project may have had an idea that was essential to the working out of the problem — even though by the time the project is completed it may have so changed from the original conception that the germinal idea is no longer central. The more interactive the process, the less we can retrospectively divide the work into parts corresponding to particular roles or contributions.

Furthermore, faculty members often rely on their own familiarity with the conventions of the discipline regarding coauthorship and other forms of credit, forgetting that students and other participants in a project may be unacquainted with them. Again, I wonder whether departments or laboratories could ameliorate the bitterness of disappointed expectations by a general discussion, in advance, of the ground rules. The understanding in my laboratory was thus: If I had contributed to the idea of the project and had also contributed significantly to the hands-on work, coauthorship was justified; but any coauthor had to have a complete enough grasp of the whole effort to defend it effectively in a scientific meeting. This test, of course, is tailored to an experimental science and surely is not the only one applicable even there. Whatever the agreement, it is necessary also that there be a prior understanding of the the scope of the particular project or sub-project — that is, both should know the anticipated product to which the agreement applies.

Third: Another aspect of the same cluster of issues — who may publish first, who must consent, what connections with the work need be acknowledged and how — is associated particularly with review articles, books (or chapters of books), or symposium contributions, especially "state of the discipline" pieces. Opportunities to produce more comprehensive works of this kind come mainly to senior scholars. In describing the significant developments in one's field, there is a natural tendency to include work done by oneself and one's students and junior associates. In the usual case, the scope of the topic is broad enough so that including all associates as coauthors is impractical and / or silly (although in a few cases it may be managed). Where the piece deals with data or results of others that are already published as a paper or dissertation, or have been accepted for publication, then employing them with appropriate citation is obviously proper.

If the material is yet unpublished but will be issued as a joint work, I think it is generally accepted that any of its prospective coauthors may refer to it, even at length, in a

separate work of sole authorship — provided that its joint origin is prominently acknowledged and provided that the opportunity for regular scholarly publication is not preempted. Common courtesy as well as an appropriate concern for the welfare of coauthors of the yet-unpublished work — especially, of course, if they are one's own students — requires that they be consulted and that reasonable requests be accommodated. If the material is as yet unpublished and if it will not be a joint work, permission must be obtained for any extended discussion and should be obtained as a matter of courtesy even for a passing notice. (Acknowledgment of the source is always essential, of course; if one permits it to be understood that the work is one's own, I think we are back to plagiarism of the simple kind.)

Fourth: There is a tight coupling between authorship and responsibility. Let us suppose that the name of a faculty member has been included on a paper resulting from the relatively independent experiments done by a student or fellow. If the data are then shown to be faulty, or worse, invented, it seems clear to me that the faculty member is responsible. Indeed, as Provost Albert Hastorf's 1984 memorandum on academic fraud pointed out, faculty members are generally responsible for the scholarly conduct of staff and students involved in their research enterprises. When one assumes coauthorship, a still higher duty of certainty prevails. The defense of minimal participation in work done in one's laboratory is generally questionable; surely it is entirely inapplicable when one is coauthor of the disputed work.

Once again, however, clarity is clouded by differences in disciplinary culture. An individual's place in the list of authors of a work may or may not be a meaningful signal about the degree of contribution: in some fields the authorship sequence is rich in meaning, elsewhere it may be entirely empty. Even within a discipline, customs vary: In most biological papers, the sequence of authors is in approximate order of extent of contribution; but there is also a tradition that places the chief of the laboratory last regardless of relative weight of contribution, and at least one leading journal extracts all meaning from sequence by requiring alphabetical listing of authors.

But those complexities only underscore the importance of establishing sound principles for determining coauthorship and provide further reason for extreme care by faculty members. While I understand the need to respect the nuances that readers (including prospective employers) will derive from author sequence in a journal article, it does seem to me that as a prima facie matter, each coauthor of a work is accountable for its authenticity and quality. Shared credit should entail shared responsibility. Where, by custom or agreement, that is not the case — in fields where 10 or more names regularly appear on a single paper as coauthors, for example, or where the work is a collaboration between scholars in different fields who lack intimate command of one another's areas — some explicit disclosure of that fact would seem desirable.

Finally, let me offer an observation on a different topic, one that has an oblique relationship to the issues considered above. I note that the pressures have increased — unacceptably, in some disciplines — to produce immense lists of publications. Those who have served on the Advisory Board, or who have seen promotion and tenure files from throughout the University for other reasons, will know the extraordinary range in numbers of publications. In the humanities, most social sciences, and some natural sciences, only a handful of works will be listed. In some Medical School departments especially, there may be more than 100 publications in a period of five or six years.

Again, the functions and traditions of publication vary by field: In some, it is (or seems) important that each small step be documented and circulated as soon as possible. In other disciplines, even primary research work is published only after extensive collection of data and the attainment of some theoretical outcome. Works of the same size can, of course, be more or less significant, as scholars in the field will immediately discern.

Nonetheless, I think the exaggerated growth of publications in some fields has become pathological, and I would like to see the problem addressed here at Stanford and, perhaps, in conference with our colleagues elsewhere. It seems likely that the pressure to have a list of publications several pages long to achieve tenure, or to produce a dozen or more papers a year to maintain one's status as a productive scholar, may lead to the kinds of behavior mentioned above: to the seizing of authorial credit where it is not due and, perhaps, to ungenerous attempts to exclude others from authorship. Indeed, there are suggestions that it contributes to the likelihood of outright academic fraud. It is very well to say that such behavior is improper even in the face of great pressure. We do say it, and mean it. But should we not question the propriety of letting the pressure continue to increase, and address ourselves to means of relieving it while maintaining (indeed, enhancing) the quality of scholarly output?

These reflections do not cover the full range of problems in connection with authorship, acknowledgment, and use of intellectual product. As I said, I hope they will provoke discussion and, if not resolution, at least steps toward the expression of commonly held principles. Surely many problems can be ameliorated by a constant awareness of the special duty of faculty to fostering the intellectual growth and independence of their students, by a habit of generosity in giving as much credit to their contributions as is consistent with realistic appraisal, and by the meticulous observation of strict standards of citation and acknowledgment. Others are harder: they may ultimately yield, if at all, only to the careful development of consensus about the ethics of our profession.

To begin our consideration of these issues, I am asking the Committee on Research to examine them during the coming academic year and to forward their views to the provost and to me. We shall then discuss their further disposition with the Academic Senate.